GEORGIA INSTITUTE OF TECHNOLOGY
ENGINEERING EXPERIMENT STATION

PROJECT NO. A-398

-FINAL REPORT-

SATELLITE ORBIT DETERMINATION AND ERROR ANALYSES OF PROCEDURES

By

J. C. Currie
E. L. Davis, Jr.
J. R. Garrett
I. E. Perlin
J. W. Wray

Contract No. AF 29(600)-1756
Department of the Air Force

Project No. 1770
Task No. 17703

September, 1959

Headquarters
Air Force Missile Development Center
Air Research and Development Command
United States Air Force
Holloman Air Force Base, New Mexico

ABSTRACT

The first section develops a method using Chebyshev polynomials for the multi-precision computation of Arctan x, Arcsin x, and Arccos x. An error analysis for computers using floating point arithmetic is made.

The second section gives a variation on that of Section 1. A comparison of the two techniques regarding errors and computer operations involved is made.

The third section initially studies the well-known methods of orbit determination for earth satellites using electronic computers. A discussion of truncation, round-off and propagated error is given, and a generalized procedure for evaluating these errors is presented. The procedure is applied to Cowell's Method of orbit calculation. The truncation error is found to be of the order of h^{2n+1} where n is the order of differences used and h is the time step in the numerical integration. The round-off error is expressed in statistical terms. The machine program used in this investigation is given in some mathematical detail. The general recommendations are that the problem be formulated in rectangular coordinates, solved using Cowell's Method, and differential corrections be applied to computed results.

The fourth section gives formulas for differentially correcting the orbit of a near earth satellite. The formulas are based on an approximate solution, in rectangular coordinates, of the differential equations of motion. Drag and oblateness perturbations are included. The rectangular coordinates of velocity and position at a reference time are adjusted.

In addition to the authors, the following staff members participated in the work described in this report: J. G. Stein and C. P. Reed, Jr.

Wm. F. Atchison
Project Director

J. E. Boyd, Director
Engineering Experiment Station

TABLE OF CONTENTS

TABLE OF CONTENTS

* * *

FOREWORD

The essence of the work contained herein appeared initially as four separate interim technical reports as the study progressed. The first two sections were written by I. E. Perlin. The third section represents the efforts in varying degrees of all five authors. The fourth section was written by E. L. Davis, Jr., and J. R. Garrett.

I. HIGH PRECISION CALCULATION OF ARCSIN X, ARCCOS X, AND ARCTAN X[1]

1. Introduction

Since Arcsin x and Arccos x can be expressed in terms of the Arctangent by means of the identities

$$\text{Arcsin } x = \text{Arctan } \frac{x}{\sqrt{1-x^2}} \quad ,$$

$$\text{Arccos } x = \frac{\pi}{2} - \text{Arctan } \frac{x}{\sqrt{1-x^2}} \quad ,$$

a single subroutine, that for the Arctangent, will suffice to compute the three inverse trigonometric functions Arcsin x, Arccos x, and Arctan x.

In the second part of the paper, an approximating polynomial for Arctan x will be developed. In order to keep the truncation error as small as desired, it is necessary to choose the degree of the approximating polynomial and the interval over which the approximation is valid. This choice will be made in the third part of this paper. In the fourth part of this paper, a method will be given for the calculation of the Arctan x for the values of x outside the interval selected in the third part. In the fifth section an error analysis will be given.

2. Polynomial Approximation for Arctan x

The expansion

$$(2.1) \qquad \text{Arctan } (x \tan 2\theta) = 2 \sum_{i=o}^{\infty} \frac{(-1)^i (\tan \theta)^{2i+1}}{2i+1} \quad T_{2i+1} (x) \ ,$$

where $T_i(x)$ are the Chebyshev polynomials, i.e.

$$T_i (\cos n) = \cos (in) \quad ,$$

will be used to obtain an approximating polynomial for the Arctangent.

To establish (2.1), expand

$$f(y) = \text{Arctan } \frac{2a \cos y}{1-a^2}$$

into a Fourier Cosine series in the interval $0 \leq y \leq \pi$. The parameter a is chosen such that $0 < a < 1$.

Then,

$$f(y) = \frac{1}{2} \ A_o(a) + \sum_{j=1}^{\infty} \ A_j(a) \ \cos (jy) \quad ,$$

[1]Released by the authors. September 1959

where

$$\pi A_j(a) = 2 \int_0^\pi f(y) \cos(jy)\, dy \qquad (j=1, 2, \ldots) \; .$$

Integrating by parts,

$$\pi j A_j(a) = 2a(1-a^2) \int_0^\pi \frac{\cos([j-1]y) - \cos([j+1]y)}{1+2a^2 \cos 2y + a^4} \, dy \; .$$

Now,[1]

$$\frac{1-a^4}{1+2a^2 \cos(2y) + a^4} = 1 + 2 \sum_{k=1}^{\infty} (-1)^k a^{2k} \cos(2ky) \; .$$

Hence,

$$\int_0^\pi \frac{\cos([2j-1]y)}{1+2a^2 \cos(2y) + a^4} \, dy = 0 \qquad (j=1, 2, \ldots) \; ,$$

and

$$A_{2j} = 0 \qquad (j=1, 2, \ldots) \; .$$

However,

$$\pi (2j+1) A_{2j+1}(a) = \frac{4a}{1+a^2} \int_0^\pi \left[(-1)^j a^{2j} \cos^2(2jy) - (-1)^{j+1} a^{2j+2} \cos^2([2j+2]y) \right] dy \; ,$$

$$\pi (2j+1) A_{2j+1}(a) = 2a \pi (-1)^j a^{2j}$$

and

$$A_{2j+1}(a) = \frac{2(-1)^j a^{2j+1}}{2j+1}$$

The expansion (2.1) has now been obtained. This expansion is absolutely and uniformly convergent for $|x| \leq 1$ and $0 \leq \theta < \frac{\pi}{4}$.

[1]See A. Erdélyi, Higher Transcendental Functions, Vol.2,p.186, McGraw-Hill Book Company (1953)

The convergence can readily be seen from the fact that

$$\left| \frac{(-1)^i (\tan \theta)^{2i+1} T_{2i+1}(x)}{2i+1} \right| \leqq \left| (\tan \theta)^{2i+1} \right| ,$$

since $|T_i(x)| \leqq 1$, and the dominating series is a geometric series with common ratio $\tan^2 \theta < 1$.

To obtain an approximating polynomial, the expansion (2.1) is truncated after n terms. Thus, the approximating polynomial is

$$P_{n-1} (x \tan 2\theta) = 2 \sum_{i=o}^{n-1} (-1)^i \frac{(\tan \theta)^{2i+1} T_{2i+1}(x)}{2i+1} .$$

The truncation error is

$$R_n = 2 \sum_{i=n}^{\infty} \frac{(-1)^i (\tan \theta)^{2i+1} T_{2i+1}(x)}{2i+1} .$$

Now,

$$\left| R_n \right| \leqq 2 \sum_{i=n}^{\infty} \left| (\tan \theta)^{2i+1} \right| ,$$

since,

$$\left| \frac{T_{2i+1}(x)}{2i+1} \right| \leqq \left| x \right| .$$

And

$$(2.2) \qquad \left| R_n \right| \leqq 2 \tan \theta \frac{(\tan \theta)^{2n}}{1-\tan^2\theta} \left| x \right| = \tan 2 \theta (\tan \theta)^{2n} \left| x \right| .$$

The quantities n and $\tan \theta$ can now be selected to make the truncation error as small as desired.

To obtain the polynomial in $M = x \tan 2\theta$, the Chebyshev polynomials, $T_{2i+1}(x)$, are expressed in terms of x by[2]

$$T_{2i+1}(x) = (-1)^i \frac{2i+1}{2} \sum_{r=0}^{i} (-1)^r \frac{\binom{i+r}{i-r} (2x)^{2r+1}}{2r+1} ,$$

and 2x is replaced by $\frac{M(1-\tan^2\theta)}{\tan\theta}$.

Then

$$P_{n-1}(M) = \sum_{r=0}^{n-1} \frac{(-1)^r B_{nr} M^{2r+1}}{2r+1} , \tag{2.3}$$

where

$$B_{nr} = (1-\tan^2\theta)^{2r+1} \sum_{k=0}^{n-r-1} \binom{2r+k}{k} (\tan\theta)^{2k} .$$

3. Selection of the Degree and Interval

To insure that the approximating polynomial (2.3) is accurate to twenty decimal places, n and $\tan\theta$ must be chosen so that

$$|R_n| < 5 \ (10)^{-21} .$$

It can be seen from the expression for R_n given by (2.2) that this can be accomplished by choosing $n = 9$ and $\tan\theta = \tan\frac{\pi}{48}$.

The approximating polynomial

$$P_8(M) = \sum_{r=0}^{8} \frac{(-1)^r C_{9r} M^{2r+1}}{2r+1} , \tag{3.1}$$

where

$$C_{9r} = (1-\tan^2\tfrac{\pi}{48})^{2r+1} \sum_{k=0}^{8-r} \binom{2r+k}{k} (\tan\tfrac{\pi}{48})^{2k} ,$$

will yield the Arctan M accurate to twenty decimal places (actually twenty-one decimal places) in the range $|M| \leq \tan\frac{\pi}{24}$.

[2]See A. Erdélyi, Higher Transcendental Functions, Vol.2, p.185, McGraw-Hill Book Company (1953)

4. Calculation of Arctan M

Subdivide the range 0 to ∞ into seven intervals as follows:

$$I_1 \; : \; 0 \leqq u < \tan \frac{\pi}{24} \quad ,$$

$$I_2 \; : \; \tan \frac{\pi}{24} \leqq u < \tan \frac{3\pi}{24} \quad ,$$

$$\cdots$$

$$I_j \; : \; \tan\frac{(2j-3)\pi}{24} \leqq u < \tan\frac{(2j-1)\pi}{24} \quad ,$$

$$\cdots$$

$$I_7 \; : \; \tan \frac{11\pi}{24} \leqq u < \tan \frac{\pi}{2} \quad .$$

If $|M| \subset I_1$, use the approximating polynomial (3.1) to calculate Arctan $|M|$. If $|M| \subset I_{j+1}$ ($j = 1,2,\ldots,5$), then the addition formula

$$\text{(4.1)} \qquad \text{Arctan } |M| = \frac{j\pi}{12} + \text{Arctan } t_j \quad ,$$

where

$$t_j = \frac{|M| - \tan \frac{j\pi}{12}}{1 + |M| \tan \frac{j\pi}{12}}$$

can be used to obtain a value of t_j such that

$$|t_j| \leqq \tan \frac{\pi}{24} \quad .$$

The value of Arctan $|t_j|$ can then be calculated from (3.1) and Arctan $|M|$ computed by (4.1).

If $|M| \subset I_7$, then

$$\text{(4.2)} \qquad \text{Arctan } |M| = \frac{\pi}{2} - \text{Arctan} \frac{1}{|M|} \quad ,$$

and

$$0 < \frac{1}{|M|} < \tan \frac{\pi}{24} \quad .$$

In this case, Arctan $\frac{1}{|M|}$ can be calculated by (3.1), and then Arctan $|M|$ by (4.2).

5. Error Analysis

A. General

It shall be assumed that all positive numbers, x, are represented in the computer as

$$x = (x_1 \beta^{-1} + x_2 \beta^{-2} + \ldots + x_\lambda \beta^{-\lambda}) \beta^\rho \quad , \qquad (0 \leq x_i < \beta) \tag{5.1}$$

where $x_1 \neq 0$, ρ is an integer, and β is either the base 2 or 10. Numbers in the form (5.1) shall be called computer numbers.

The errors arising in computations by a computer can be classified into three categories:

a) Truncation errors that arise from the approximation of an infinite process by a finite one.

b) Propagation errors that arise from the approximation of a number by another number.

c) Round-off errors that arise from the approximation of the sum, difference, product, or quotient of two computer numbers by a computer number. These errors will be treated as indicated below.

1. Truncation error:

Bounds for the truncation error are obtained by determining dominants for the remainder after a finite number of terms.

2. Propagation error:

If x and y are approximated by x' and y' respectively, and the errors in each are denoted by $\epsilon(x)$ and $\epsilon(y)$, then the propagated error in the sum or difference, product, and quotient satisfy the inequalities

$$|\epsilon(x \pm y)| \leq |\epsilon(x)| + |\epsilon(y)| \quad ,$$

$$\left|\frac{\epsilon(x\,y)}{x\,y}\right| \leq \left|\frac{\epsilon(x)}{x}\right| + \left|\frac{\epsilon(y)}{y}\right| \qquad (xy \neq 0) \quad ,$$

$$\left|\frac{\epsilon(\frac{x}{y})}{\frac{x}{y}}\right| \leq \frac{\left|\frac{\epsilon(x)}{x}\right| + \left|\frac{\epsilon(y)}{y}\right|}{1 - \left|\frac{\epsilon(y)}{y}\right|} \qquad (xy \neq 0) \quad .$$

It is assumed, of course, that the relative errors in x and y are small.

If $f(x)$ represents an analytic function of x, then the propagated error in $f(x)$ due to an error in x satisfies the inequality

$$\left| \epsilon \left[f(x) \right] \right| \leqq \left| f'(x) \right| \left| \epsilon (x) \right| \quad ,$$

provided $f'(x) \neq 0$, and it is assumed that the error in x is small so that higher order terms can be neglected.

3. Round-off error:

If x and y are computer numbers, and $x \circ y$ represents any of the four elementary operations, addition, subtraction, multiplication, or division, then the round-off error in $x \circ y$, designated by $\epsilon(x \circ y)$, can be bounded as is shown.

Let

$$x = (x_1 \beta^{-1} + x_2 \beta^{-2} + \ldots + x_\lambda \beta^{-\lambda}) \beta^{\rho} \qquad (x_1 \neq 0) \quad ,$$

and

$$y = (y_1 \beta^{-1} + y_2 \beta^{-2} + \ldots + y \beta^{-\lambda}) \beta^{\sigma} \qquad (y_1 \neq 0) \quad .$$

Then

$$x \circ y = (z_1 \beta^{-1} + z_2 \beta^{-2} + \ldots + z_\lambda \beta^{-\lambda} + \ldots) \beta^{\tau} \quad (z_1 \neq 0) \quad .$$

The computer number representing $x \circ y$ is obtained by truncating $x \circ y$ after the term $z_\lambda \beta^{-\lambda}$. It shall be assumed that this truncation is accomplished with rounding, i.e., if $z_{\lambda+1} \geqq \frac{1}{2}\beta$, then add $\beta^{-\lambda}$ to $z_1 \beta^{-1} + z_2 \beta^{-2} + \ldots + z_\lambda \beta^{-\lambda}$, and drop all terms beyond this point. If $z_{\lambda+1} < \frac{1}{2}\beta$, drop all terms beyond that one involving $\beta^{-\lambda}$. The round-off error in representing $x \circ y$ by a computer number designated by $\overline{x \circ y}$, has the bound

$$\left| \epsilon(x \circ y) \right| \leqq \left| \overline{x \circ y} \right| \left(\tfrac{1}{2}\beta \right) \beta^{-\lambda}$$

B. Analysis of error in the calculation of Arctan x

1. Suppose $0 \leqq x < \tan \frac{\pi}{24}$.

a. Consider the truncation error.

If $0 \leqq x < \tan \frac{\pi}{24}$, then the truncation error in calculating Arctan x by (3.1) is given by (2.2) in the form

$$(5.2) \qquad | \epsilon_T | \leqq x \left(\tan \tfrac{\pi}{48}\right)^{18} < (6)\, 10^{-22}\, x < 2^{-70}\, x \quad .$$

b. Consider the propagated error due to errors in the coefficients of the approximating polynomial.

The coefficients in the approximating polynomial (3.1) are bounded by the integer 1. This fact is demonstrated as follows:

$$C_{90} = 1 - \tan^2 \tfrac{\pi}{48} < 1 \quad .$$

$$\frac{C_{9r}}{2r+1} = \left(\frac{1 - \tan^2 \frac{\pi}{48}}{2r+1}\right)^{2r+1} \sum_{k=o}^{8-r} \binom{2r+k}{k} \left(\tan \tfrac{\pi}{48}\right)^{2k} \qquad (r = 1,2,\ldots,8) \ ,$$

Now

$$\frac{\binom{2r+k+1}{k+1} \left(\tan \frac{\pi}{48}\right)^{2k+2}}{\binom{2r+k}{k} \left(\tan \frac{\pi}{48}\right)^{2k}} = \left(\frac{2r+k+1}{k+1}\right) \tan^2 \tfrac{\pi}{48} \quad ,$$

$$< 17 \quad \tan^2 \tfrac{\pi}{48} \quad ,$$

$$< \quad 0.09 \quad .$$

Therefore,

$$\frac{C_{9r}}{2r+1} < \left(\frac{1 - \tan^2 \frac{\pi}{48}}{2r+1}\right)^{2r+1} \sum_{k=o}^{8-r} (0.09)^k \quad , \qquad (r = 1,2,\ldots,8) \quad .$$

$$\frac{C_{9r}}{2r+1} < \left(\frac{1 - \tan^2 \frac{\pi}{48}}{2r+1}\right)^{2r+1} \left(\frac{1}{1 - 0.09}\right) \quad , \qquad (r = 1,2,\ldots,8) \quad .$$

$$\frac{C_{9r}}{2r+1} < \left(1 - \tan^2 \frac{\pi}{48}\right)^{2r+1} \left(\frac{1.1}{3}\right) \qquad , \qquad (r = 1,2,\ldots,8) \quad .$$

$$\frac{C_{9r}}{2r+1} < 1 \quad , \quad (r = 1,2,\ldots,8) \quad .$$

The error in P(x) due to errors in the coefficients is bounded as follows:

$$\left| \epsilon[P(x)] \right| \leqq (\tfrac{1}{2}\beta)\ (\beta^{-\lambda})\ (x + x^3 + \ldots + x^{17}) \quad ,$$

$$\left| \epsilon[P(x)] \right| < (\tfrac{1}{2}\beta)\ (\beta^{-\lambda})\ \frac{x}{1 - x^2} \quad ,$$

$$(5.3) \quad \left| \epsilon[P(x)] \right| < (1.02)\ (\tfrac{1}{2}\beta)\ (\beta^{-\lambda})\ x \quad .$$

c. Consider the Round-off error.

The polynomial P(x) will be calculated as follows:

$$p_1 = a_{15} + a_{17}\, x^2 \quad ,$$
$$p_2 = a_{13} + p_1\, x^2 \quad ,$$
$$p_3 = a_{11} + p_2\, x^2 \quad ,$$
$$\ldots$$
$$p_8 = a_1 + p_7\, x^2 \quad ,$$
$$P(x) = p_8\, x \quad .$$

Errors are caused by round-off errors in multiplication and addition and by propagated errors. It is seen that

$$|\epsilon(p_1)| \leqq |a_{17}|\, x^2 (\tfrac{1}{2}\beta)\beta^{-\lambda} + |a_{17}\, x^2|\, (\tfrac{1}{2}\beta)\beta^{-\lambda} + |p_1|\, (\tfrac{1}{2}\beta)\beta^{-\lambda} \quad ,$$

$$|\epsilon(p_2)| \leqq |p_1|\, x^2 (\tfrac{1}{2}\beta)\beta^{-\lambda} + |\epsilon(p_1)|\, x^2 + |p_1\, x^2|\, (\tfrac{1}{2}\beta)\beta^{-\lambda} + |p_2|(\tfrac{1}{2}\beta)\beta^{-\lambda} \quad ,$$

$$|\epsilon(p_3)| \leqq |p_2|\, x^2 (\tfrac{1}{2}\beta)\beta^{-\lambda} + |\epsilon(p_2)|\, x^2 + |p_2\, x^2|\, (\tfrac{1}{2}\beta)\beta^{-\lambda} + |p_3|(\tfrac{1}{2}\beta)\beta^{-\lambda} \quad ,$$

$$\ldots$$

$$|\epsilon(p_8)| \leqq |p_7|\, x^2 \,(\tfrac{1}{2}\beta)\, \beta^{-\lambda} + |\epsilon(p_7)|\, x^2 + |p_7 x^2|(\tfrac{1}{2}\beta)\,\beta^{-\lambda} + |p_8|(\tfrac{1}{2}\beta)\,\beta^{-\lambda} \quad ,$$

$$|\epsilon[P(x)]| \leqq |\epsilon(p_8)|\, x + |p_8\, x|\,(\tfrac{1}{2}\beta)\,\beta^{-\lambda} \quad .$$

Since $|a_i| \leqq 1$ and $|p_i| \leqq 1$,

$$|\epsilon(p_1)| \leqq 2\, x^2\, (\tfrac{1}{2}\beta)\,\beta^{-\lambda} + (\tfrac{1}{2}\beta)\,\beta^{-\lambda} \quad ,$$

$$|\epsilon(p_2)| \leqq 2\, x^2\, (\tfrac{1}{2}\beta)\,\beta^{-\lambda} + |\epsilon(p_1)|\, x^2 + (\tfrac{1}{2}\beta)\,\beta^{-\lambda} \quad ,$$

$$\cdots$$

$$|\epsilon(p_8)| \leqq 2\, x^2\, (\tfrac{1}{2}\beta)\,\beta^{-\lambda} + |\epsilon(p_7)|\, x^2 + (\tfrac{1}{2}\beta)\,\beta^{-\lambda} \quad ,$$

$$|\epsilon[P(x)]| \leqq x\, (\tfrac{1}{2}\beta)\,\beta^{-\lambda} + |\epsilon(p_8)|\, x \quad .$$

Multiplying the first inequality by x^{15}, the second by x^{13}, etc., and adding the following inequality is obtained.

$$|\epsilon[P(x)]| \leqq 2\,(\tfrac{1}{2}\beta)\,\beta^{-\lambda}\,(x + x^3 + \ldots + x^{17}) + (\tfrac{1}{2}\beta)\beta^{-\lambda}\,(x^3 + x^5 + \ldots + x^{15}),$$

Hence,

$$|\epsilon[P(x)]| < \frac{2 + x^2}{1 - x^2}\,(\tfrac{1}{2}\beta)\,\beta^{-\lambda}\, x \quad ,$$

(5.4) $$|\epsilon[P(x)]| < (2.06)\,(\tfrac{1}{2}\beta)\beta^{-\lambda}\, x \quad .$$

By combining (5.2), (5.3), and (5.4), the total error in Arctan x is bounded by

(5.5) $$|\epsilon(\text{Arctan } x)| < (3.205)\,(\tfrac{1}{2}\beta)\beta^{-\lambda}\, x \quad ,$$

if $\beta = 10$ and $\lambda \leqq 21$, or if $\beta = 2$ and $\lambda \leqq 67$.

Finally, the relative error in Arctan x can be bounded as follows:

$$\left|\frac{\epsilon(\text{Arctan } x)}{\text{Arctan } x}\right| < \frac{x}{\text{Arctan } x}\left|\frac{\epsilon(\text{Arctan } x)}{x}\right| \quad ,$$

$$\left|\frac{\epsilon(\text{Arctan } x)}{\text{Arctan } x}\right| < (1.006)\left|\frac{\epsilon(\text{Arctan } x)}{x}\right| \quad ,$$

since

$$1 < \frac{x}{\text{Arctan } x} < \frac{1}{1 - \frac{1}{3}x^2} < 1.006 \quad .$$

Hence, if $\beta = 2$ or 10,

$$(5.6) \qquad \left| \frac{\epsilon\,(\text{Arctan } x)}{\text{Arctan } x} \right| < (3.3) \quad (\tfrac{1}{2}\ \beta)\ \beta^{-\lambda} \qquad .$$

If $\beta = 10$, then

$$\left| \frac{\epsilon\,(\text{Arctan } x)}{\text{Arctan } x} \right| < (1.7) \quad 10^{-\lambda+1} \qquad ,$$

and the error is at most $\pm$ 2 in the (λ -1)-st significant digit.

If $\beta = 2$, then

$$\left| \frac{\epsilon\,(\text{Arctan } x)}{\text{Arctan } x} \right| < \beta^{-\lambda+2}$$

and λ -3 significant digits are correct.

2. Suppose $\tan \frac{\pi}{24} \leqq x < \tan \frac{11\pi}{24}$.

The calculation of Arctan x for x in this range is made by means of

$$\text{Arctan } x = \frac{j\pi}{12} + \text{Arctan } t_j \qquad ,$$

where

$$t_j = \frac{x - \tan \frac{j\pi}{12}}{1 + x \tan \frac{j\pi}{12}} \qquad .$$

The number $\tan \frac{j\pi}{12}$ as stored in the computer is regarded as exact, but the angle $\frac{j\pi}{12}$ will be in error. The error in $\frac{j\pi}{12}$ will be taken into consideration.

Let N denote the numerator of the fraction yielding t_j, and D the denominator. Then

$$\left| \frac{\epsilon(N)}{N} \right| \leqq \quad (\tfrac{1}{2}\,\beta\ )\ \beta^{-\lambda} \qquad ,$$

and

$$\left| \frac{\epsilon(D)}{D} \right| \leqq 2\ (\tfrac{1}{2}\,\beta\ )\ \beta^{-\lambda} \qquad .$$

The error in t_j is

$$\left| \frac{\epsilon(t_j)}{t_j} \right| \leq \frac{3\left(\frac{1}{2}\beta\right)\beta^{-\lambda}}{1-\beta^{-\lambda+1}} + \left(\frac{1}{2}\beta\right)\beta^{-\lambda} \quad ,$$

$$\left| \frac{\epsilon(t_j)}{t_j} \right| < \left[3\left(1+2\beta^{-\lambda+1}\right) + 1 \right] \left(\frac{1}{2}\beta\right)\beta^{-\lambda} \quad .$$

The propagated error in Arctan t_j due to this error in t_j is

$$\left| \epsilon(\text{Arctan } t_j) \right| \leq \frac{1}{1+t_j^{\,2}} \, t_j \left[3\left(1+2\beta^{-\lambda+1}\right) + 1 \right] \left(\frac{1}{2}\beta\right)\beta^{-\lambda} \quad ,$$

$$\left| \frac{\epsilon(\text{Arctan } t_j)}{\text{Arctan } t_j} \right| \leq \frac{t_j}{(1+t_j^{\,2})\text{Arctan } t_j} \left| 3\left(1+2\beta^{-\lambda+1}\right) + 1 \right| \left(\frac{1}{2}\beta\right)\beta^{-\lambda} \quad ,$$

$$\left| \frac{\epsilon(\text{Arctan } t_j)}{\text{Arctan } t_j} \right| \leq \left| 3\left(1+2\beta^{-\lambda+1}\right) + 1 \right| \left(\frac{1}{2}\beta\right)\beta^{-\lambda} \quad ,$$

since[3]

$$\frac{A}{(1+A^2)\text{Arctan } A} \leq 1 \quad .$$

The total error in calculating Arctan t_j is, then,

$$\left| \epsilon(\text{Arctan } t_j) \right| \leq \left| \text{Arctan } t_j \right| \left\{ \left[3\left(1+2\beta^{-\lambda+1}\right) + 1 \right] + 3.3 \right\} \left(\frac{1}{2}\beta\right)\beta^{-\lambda} \quad ,$$

$$\left| \epsilon(\text{Arctan } t_j) \right| \leq \left| \text{Arctan } t_j \right| \left(7.3 + 6\beta^{-\lambda+1}\right) \left(\frac{1}{2}\beta\right)\beta^{-\lambda} \quad .$$

[3] See E. P. Adams and R. L. Hippisley, Smithsonian Mathematical Formulae and Tables of Elliptic Functions, p.122, Smithsonian Institution, Washington, D. C. (1922).

Now, the error in taking $\frac{j\pi}{12}$ as the angle in computing Arctan x does not exceed $\frac{j\pi}{12}\left(\frac{1}{2}\beta\right)\beta^{-\lambda}$. This is seen from the following. An error in x will result in a propagated error in Arctan x which can be determined by

$$\left|\frac{\epsilon(\text{Arctan } x)}{\text{Arctan } x}\right| \leqq \frac{x}{(1+x^2)\text{Arctan } x}\left|\frac{\epsilon(x)}{x}\right| ,$$

$$\left|\frac{\epsilon(\text{Arctan } x)}{\text{Arctan } x}\right| \leqq \left|\frac{\epsilon(x)}{x}\right| .$$

Hence, taking $\frac{j\pi}{12}$ as the angle whose tangent is $\tan\frac{j\pi}{12}$ as stored in the computer results in an error of at most $\frac{j\pi}{12}\left(\frac{1}{2}\beta\right)\beta^{-\lambda}$.

To continue with the error analysis, it must be remembered that $\frac{\pi}{2}$ is stored and $\frac{j\pi}{12}$ computed. Then

$$\begin{aligned}\left|\epsilon(\text{Arctan } x)\right| \leqq\ & \frac{j\pi}{12}\left(\frac{1}{2}\beta\right)\beta^{-\lambda} + \frac{j\pi}{12}\left(\frac{1}{2}\beta\right)\beta^{-\lambda} \\ & + \frac{j\pi}{12}\left(\frac{1}{2}\beta\right)\beta^{-\lambda} \\ & + \text{Arctan } x\left(\frac{1}{2}\beta\right)\beta^{-\lambda} \\ & + \left|\text{Arctan } t_j\right|\left(7.3 + 6\beta^{-\lambda+1}\right)\left(\frac{1}{2}\beta\right)\beta^{-\lambda} ,\end{aligned}$$

$$\left|\frac{\epsilon(\text{Arctan } x)}{\text{Arctan } x}\right| \leqq \left\{\frac{\frac{j\pi}{4}}{\text{Arctan } x} + 1 + \frac{\text{Arctan } t_j}{\text{Arctan } x}\left(7.3 + 6\beta^{-\lambda+1}\right)\right\}\left(\frac{1}{2}\beta\right)\beta^{-\lambda} ,$$

$$\begin{aligned}\left|\frac{\epsilon(\text{Arctan } x)}{\text{Arctan } x}\right| &\leqq \left\{\frac{\frac{j\pi}{4}}{\frac{j\pi}{12}-\frac{\pi}{24}} + 1 + \frac{\frac{\pi}{24}}{\frac{j\pi}{12}-\frac{\pi}{24}}\left(7.3 + 6\beta^{-\lambda+1}\right)\right\}\left(\frac{1}{2}\beta\right)\beta^{-\lambda} , \\ &\leqq \left\{\frac{6j}{2j-1} + 1 + \frac{7.3 + 6\beta^{-\lambda+1}}{2j-1}\right\}\left(\frac{1}{2}\beta\right)\beta^{-\lambda} , \\ &\leqq \frac{8j + 6.3 + 6\beta^{-\lambda+1}}{2j-1}\left(\frac{1}{2}\beta\right)\beta^{-\lambda} , \\ &\leqq \left(14.3 + 6\beta^{-\lambda+1}\right)\left(\frac{1}{2}\beta\right)\beta^{-\lambda} .\end{aligned}$$

If $\beta = 10$ and $\lambda \leqq 21$, then $\left|\frac{\epsilon\ (\text{Arctan}\ x)}{\text{Arctan}\ x}\right| \leqq (7.2)\ 10^{-\lambda+1}$,

and the error is at most $\pm$ 8 in the $(\lambda -1)$-st significant digit.

If $\beta = 2$ and $\lambda \leqq 67$, then $\left|\frac{\epsilon\ (\text{Arctan}\ x)}{\text{Arctan}\ x}\right| < (1.8)\ 2^{-\lambda+3}$,

and $\lambda -5$ significant digits are correct.

3. $\tan \frac{11\pi}{24} \leqq x < \infty$.

In this case Arctan x is computed by

$$\text{Arctan}\ x = \frac{\pi}{2} - \text{Arctan}\ \frac{1}{x} \quad .$$

Now,

$$\left|\epsilon\left(\frac{1}{x}\right)\right| \leqq \frac{1}{x}\ \left(\frac{1}{2}\ \beta\right)\ \beta^{-\lambda} \quad ,$$

and the propagated error in Arctan $\frac{1}{x}$ due to this error in $\frac{1}{x}$ is

$$\left|\frac{\epsilon\ (\text{Arctan}\ \frac{1}{x})}{\text{Arctan}\ \frac{1}{x}}\right| \leqq \left(\frac{1}{2}\ \beta\ \right)\beta^{-\lambda} \quad .$$

The total relative error in Arctan $\frac{1}{x}$ is

$$\left|\frac{\epsilon\ (\text{Arctan}\ \frac{1}{x})}{\text{Arctan}\ \frac{1}{x}}\right| \leqq \left(\frac{1}{2}\ \beta\right)\beta^{-\lambda} + (3.3)\ \left(\frac{1}{2}\ \beta\ \right)\beta^{-\lambda} \quad ,$$

$$\left|\frac{\epsilon\ (\text{Arctan}\ \frac{1}{x}}{\text{Arctan}\ \frac{1}{x}}\right| \leqq (4.3)\ \ \left(\frac{1}{2}\ \beta\right)\beta^{-\lambda} \quad .$$

The total relative error in Arctan x is

$$\left|\frac{\epsilon\ (\text{Arctan}\ x)}{\text{Arctan}\ x}\right| \leqq \left[\frac{(\text{Arctan}(\frac{1}{x}))(4.3) + \frac{\pi}{2}}{\text{Arctan}\ x} + 1\right] \left(\frac{1}{2}\ \beta\right)\beta^{-\lambda} \ ,$$

$$\left|\frac{\epsilon\ (\text{Arctan}\ x)}{\text{Arctan}\ x}\right| \leqq (2.5)\ \left(\frac{1}{2}\ \beta\right)\beta^{-\lambda} \quad .$$

If $\beta = 10$ and $\lambda \leqq 21$, then

$$\left| \frac{\epsilon\ (\text{Arctan } x)}{\text{Arctan } x} \right| \leq (12.5)\ 10^{-\lambda} \quad ,$$

and the error is at most $\pm$ 2 in the $(\lambda-1)$-st significant digit.

If $\beta = 2$ and $\lambda \leqq 67$, then

$$\left| \frac{\epsilon\ (\text{Arctan } x)}{\text{Arctan } x} \right| < (2.5)\ 2^{-\lambda}$$

and $\lambda - 3$ significant digits are correct.

C. Analysis of error in the calculation of Arcsin x and Arccos x

The Arcsin x is calculated by

$$\text{Arcsin } x = \text{Arctan } \frac{x}{\sqrt{1-x^2}} \quad .$$

The quantity $1-x^2$ shall be calculated as the product of $1-x$ and $1+x$.

Now,

$$\left| \epsilon\ (1-x) \right| \leqq (1-x)\ \left(\tfrac{1}{2}\beta\right)\ \beta^{-\lambda} \quad ,$$

$$\left| \epsilon\ (1+x) \right| \leqq (1+x)\ \left(\tfrac{1}{2}\beta\right)\ \beta^{-\lambda} \quad ,$$

and therefore

$$\left| \epsilon\ (1-x^2) \right| \leqq (1-x^2)\ (3)\ \left(\tfrac{1}{2}\beta\right)\ \beta^{-\lambda} \quad .$$

The propagated error in the square root of $1-x^2$ due to this error in $1-x^2$ is

$$\left| \epsilon\left(\sqrt{1-x^2}\right) \right| \leqq \frac{1}{2\sqrt{1-x^2}}\ (1-x^2)\ (3)\ \left(\tfrac{1}{2}\beta\right)\ \beta^{-\lambda} \quad .$$

If the square root is calculated by Newton's Method, then a round error of at most 1 in the last significant figure is possible. Hence the total relative error in calculating $\sqrt{1-x^2}$ is

$$\left| \frac{\epsilon\sqrt{1-x^2}}{\sqrt{1-x^2}} \right| \leqq (3.5)\ \left(\tfrac{1}{2}\beta\right)\ \beta^{-\lambda} \quad .$$

The error in the fraction $\frac{x}{\sqrt{1-x^2}}$ is

$$\left| \epsilon \left(\frac{x}{\sqrt{1-x^2}} \right) \right| \leqq \left| \frac{x}{\sqrt{1-x^2}} \right| \quad \frac{(3.5)\ (\frac{1}{2}\beta)\ \beta^{-\lambda}}{1 - (3.5)(\frac{1}{2}\beta)\ \beta^{-\lambda}} + (\tfrac{1}{2}\beta)\ \beta^{-\lambda} \ ,$$

$$\left| \epsilon \left(\frac{x}{\sqrt{1-x^2}} \right) \right| \leqq \left| \frac{x}{\sqrt{1-x^2}} \right| \left| (3.5)(1+3.5\beta^{-\lambda+1}) + 1 \right| (\tfrac{1}{2}\beta)\ \beta^{-\lambda} \ .$$

The propagated error in Arctan $\frac{x}{\sqrt{1-x^2}}$ due to this error is

$$\left| \frac{\epsilon\ (\text{Arctan}\ \frac{x}{\sqrt{1-x^2}})}{\text{Arctan}\ \frac{x}{\sqrt{1-x^2}}} \right| \leqq (4.5 + 12.5\beta^{-\lambda+1})\ (\tfrac{1}{2}\beta)\ \beta^{-\lambda} \ .$$

The total error in the computed value of Arctan $\frac{x}{\sqrt{1-x^2}}$ is then

$$\left| \frac{\epsilon\ (\text{Arctan}\ \frac{x}{\sqrt{1-x^2}})}{\text{Arctan}\ \frac{x}{\sqrt{1-x^2}}} \right| \leqq (17.8 + 18.5\beta^{-\lambda+1})\ (\tfrac{1}{2}\beta)\ \beta^{-\lambda} \ .$$

If $\beta = 10$ and $\lambda \leqq 21$, then

$$\left| \frac{\epsilon\ (\text{Arctan}\ \frac{x}{\sqrt{1-x^2}})}{\text{Arctan}\ \frac{x}{\sqrt{1-x^2}}} \right| < (10)^{-\lambda+2}$$

and the error is at most $\pm$ 1 in the $(\lambda - 2)$nd significant digit.

If $\beta = 2$ and $\lambda = 67$, then

$$\left| \frac{\epsilon\ (\text{Arctan}\ \frac{x}{\sqrt{1-x^2}})}{\text{Arctan}\ \frac{x}{\sqrt{1-x^2}}} \right| < 1.2\ (2)^{-\lambda+4} \ ,$$

and $\lambda - 6$ significant digits are correct.

The error analysis in computing Arccos x by means of

$$\text{Arccos } x = \frac{\pi}{2} - \text{Arctan } \frac{x}{\sqrt{1-x^2}}$$

is the same as that for computing Arcsin x. All that is added is a round error due to the subtraction, and this does not change the conclusion.

6. The Approximating Polynomial and Stored Constants

The approximating polynomial to be used in the range $0 \leq x < \tan \frac{\pi}{24}$ for Arctan x is

(6.1) $$P(x) = a_1 x + a_3 x^3 + \dots + a_{17} x^{17}$$

where

$$
\begin{aligned}
a_1 &= 1.0 \\
a_3 &= -0.33333\ 33333\ 33333\ 33160\ 7 \\
a_5 &= 0.19999\ 99999\ 99998\ 24444\ 8 \\
a_7 &= -0.14285\ 71428\ 56331\ 30652\ 9 \\
a_9 &= 0.11111\ 11109\ 07793\ 96739\ 3 \\
a_{11} &= -0.09090\ 90609\ 63367\ 76370\ 73 \\
a_{13} &= 0.07692\ 04073\ 24915\ 40813\ 20 \\
a_{15} &= -0.06652\ 48229\ 41310\ 82779\ 05 \\
a_{17} &= 0.05467\ 21009\ 39593\ 88069\ 41
\end{aligned}
$$

The stored constants are

(6.2)
$$
\begin{aligned}
\tan \frac{\pi}{24} &= 0.13165\ 24975\ 87395\ 85347\ 2 \\
\tan \frac{3\pi}{24} &= 0.41421\ 35623\ 73095\ 04880\ 2 \\
\tan \frac{5\pi}{24} &= 0.76732\ 69879\ 78960\ 34292\ 3 \\
\tan \frac{7\pi}{24} &= 1.30322\ 53728\ 41205\ 75586\ 8 \\
\tan \frac{9\pi}{24} &= 2.41421\ 35623\ 73095\ 04880\ 2 \\
\tan \frac{11\pi}{24} &= 7.59575\ 41127\ 25150\ 44052\ 6
\end{aligned}
$$

$$
(6.3) \quad
\begin{aligned}
\tan \frac{\pi}{12} &= 0.26794\ 91924\ 31122\ 70647\ \ 3 \\
\tan \frac{2\pi}{12} &= 0.57735\ 02691\ 89625\ 76450\ \ 9 \\
\tan \frac{3\pi}{12} &= 1 \\
\tan \frac{4\pi}{12} &= 1.73205\ 08075\ 68877\ 29352\ \ 7 \\
\tan \frac{5\pi}{12} &= 3.73205\ 08075\ 68877\ 29352\ \ 7 \\
\\
\frac{\pi}{2} &= 1.57079\ 63267\ 94896\ 61923\ \ 1 \quad .
\end{aligned}
$$

All the coefficients and stored constants have been rounded to twenty-one or twenty-two decimal places.

7. Discussion

A. Operations performed.

The procedures advanced for computing Arctan x requires the storing of: 1) nine coefficients for the approximating polynomial (6.1); 2) six constants (6.2) for locating $|x|$ in the proper interval; 3) five constants (6.3) for the calculation of $t_1, t_2, \ldots, t_5$, (these values could be calculated from (6.2)), and 4) the constant $\frac{\pi}{2}$.

For the calculation of Arctan x a maximum of eleven multiplications and one division are required. For the calculation of Arcsin x and Arccos x an additional multiplication, an additional division, and a square root are required.

B. Using less than twenty-one decimal places.

If fewer than twenty-one decimal places will be used, i.e., if $\lambda < 21$ for $\beta = 10$ or $\lambda < 67$ for $\beta = 2$, then the stored constants can be rounded to the desired number of decimal places and it may be possible to neglect the term $a_{17}\, x^{17}$ in the approximating polynomial (6.1).

C. Comparison with other methods.

In calculating Arctan x, $x > 1$, it is possible to proceed as follows:

a) Compute $\frac{1}{x}$

b) Compute Arctan $\frac{1}{x}$

c) Calculate Arctan x from Arctan $x = \frac{\pi}{2}$ - Arctan $\frac{1}{x}$.

This procedure will save five stored constants, but adds a division. It is felt that it is more economical to avoid the extra division and use the five stored constants.

Methods employing rational approximations[1] do not change the number of stored constants. The number of operations (counting multiplications and divisions) may be a few less in rational approximations, but most of these operations are divisions. In the polynomial approximation, all except one of the operations are multiplications.

[1] E. G. Kogbetliantz, Computation of Arctan N for $-\infty < N < \infty$ Using an Electronic Computer, IBM Journal of Research and Development, Vol. 2 No. 1, January 1958.

II. COMPARISON AND ERROR PREDICTION OF TWO METHODS OF CALCULATING ARCTAN X

1. INTRODUCTION

This paper is an elaboration of Technical Report No. 1, Project No. A-398 and is concerned with:

1) A comparison of two methods for calculating Arctan x.
2) Error predictions for the two methods using eight and nine terms of the approximating polynomial.

2. OUTLINE OF METHOD NO. 1

In this method twelve constants, in addition to the nine coefficients of the approximating polynomial, are stored. The twelve constants are:

$\tan \frac{\pi}{24}$	=	0.13165	24975	87395	85347	2
$\tan \frac{\pi}{12}$	=	0.26794	91924	31122	70647	3
$\tan \frac{\pi}{8}$	=	0.41421	35623	73095	04880	2
$\tan \frac{\pi}{6}$	=	0.57735	02691	89625	76450	9
$\tan \frac{5\pi}{24}$	=	0.76732	69879	78960	34292	3
$\tan \frac{\pi}{4}$	=	1.00000	00000	00000	00000	0
$\tan \frac{7\pi}{24}$	=	1.30322	53728	41205	75586	8
$\tan \frac{\pi}{3}$	=	1.73205	08075	68877	29352	7
$\tan \frac{3\pi}{8}$	=	2.41421	35623	73095	04880	2
$\tan \frac{5\pi}{12}$	=	3.73205	08075	68877	29352	7
$\tan \frac{11\pi}{24}$	=	7.59575	41127	25150	44052	6
$\frac{\pi}{2}$	=	1.57079	63267	94896	61923	1

The nine coefficients of the approximating polynomial

$$P(x) = a_1 x + a_3 x^3 \dots a_{17} x^{17}$$

are:

a_1	=	1.00000	00000	00000	00000	0
a_3	=	-0.33333	33333	33333	33160	7
a_5	=	0.19999	99999	99998	24444	8
a_7	=	-0.14285	71428	56331	30652	9
a_9	=	0.11111	11109	07793	96739	3
a_{11}	=	-0.09090	90609	63367	76370	73
a_{13}	=	0.07692	04073	24915	40813	20
a_{15}	=	-0.06652	48229	41310	82779	05
a_{17}	=	0.05467	21009	39593	88069	41

The method is as follows:

1) If $0 \leq x < \tan\frac{\pi}{24}$, calculate $P(x)$. Then

$$\text{Arctan } x = P(x).$$

2) If $\tan\frac{\pi}{24} \leq x < \tan\frac{11\pi}{24}$, first determine j such that

$$\tan\frac{(2j-1)\pi}{24} \leq x < \tan\frac{(2j+1)\pi}{24}.$$

Then calculate

$$t_j = \frac{x - \tan\frac{j\pi}{12}}{1 + x\tan\frac{j\pi}{12}}.$$

Next, calculate $P(t_j)$. Then

$$\text{Arctan } x = \frac{j\pi}{12} + P(t_j).$$

3) If $\tan\frac{11\pi}{24} \leq x < \infty$, calculate $\frac{1}{x}$. Then calculate $P(\frac{1}{x})$. Then

$$\text{Arctan } x = \frac{\pi}{2} - P(\frac{1}{x}).$$

3. OUTLINE OF METHOD NO. 2

In this method only seven constants in addition to the nine coefficients of the approximating polynomial are stored. The seven constants are:

$\tan \frac{\pi}{24} = 0.13165 \quad 24975 \quad 87395 \quad 85347 \quad 2$

$\tan \frac{\pi}{12} = 0.26794 \quad 91924 \quad 31122 \quad 70647 \quad 3$

$\tan \frac{\pi}{8} = 0.41421 \quad 35623 \quad 73095 \quad 04880 \quad 2$

$\tan \frac{\pi}{6} = 0.57735 \quad 02691 \quad 89625 \quad 76450 \quad 9$

$\tan \frac{5\pi}{24} = 0.76732 \quad 69879 \quad 78960 \quad 34292 \quad 3$

$1 = 1.00000 \quad 00000 \quad 00000 \quad 00000 \quad 0$

$\frac{\pi}{2} = 1.57079 \quad 63267 \quad 94896 \quad 61923 \quad 1$

The nine coefficients of the approximating polynomial are the same.

The method is as follows:

1) Test to see if $0 \leq x < 1$, $x = 1$, or $x > 1$.

2) If $0 \leq x < 1$, proceed as in Method No. 1.

3) If $x = 1$, then

$$\text{Arctan } x = \frac{\pi}{4}$$

4) If $x > 1$, calculate $\frac{1}{x}$. Proceed as in Method No. 1 to calculate Arctan $\frac{1}{x}$. Then

$$\text{Arctan } x = \frac{\pi}{2} - \text{Arctan } \frac{1}{x}.$$

4 . COMPARISON OF METHODS AS TO NUMBER OF OPERATIONS INVOLVED

The following table gives the number of comparisons, additions, multiplications, and divisions for the two methods.

TABLE 1

	Comparisons No. 1	Comparisons No. 2	Additions No. 1	Additions No. 2	Multiplications No. 1	Multiplications No. 2	Division No. 1	Division No. 2
$0 \leq x < \tan \frac{\pi}{24}$	1	2	8	8	10	10	0	0
$\tan \frac{\pi}{24} \leq x < \tan \frac{3\pi}{24}$	2	3	11	11	11	11	1	1
$\tan \frac{3\pi}{24} \leq x < \tan \frac{5\pi}{24}$	3	4	11	11	11	11	1	1
$\tan \frac{5\pi}{24} \leq x < 1$	4	4	11	11	11	11	1	1
$x = 1$	4	2	1	0	0	0	0	0
$1 < x < \tan \frac{7\pi}{24}$	4	6	11	12	11	11	1	2
$\tan \frac{7\pi}{24} \leq x < \tan \frac{9\pi}{24}$	5	5	11	12	11	11	1	2
$\tan \frac{9\pi}{24} \leq x < \tan \frac{11\pi}{24}$	6	4	11	12	11	11	1	2
$\tan \frac{11\pi}{24} \leq x < \infty$	6	3	9	9	10	10	1	1

The following observations can be drawn from TABLE 1.

1) For $0 \leq x < 1$ and $x \geq \tan \frac{11\pi}{24}$ both methods are identical, except that Method No. 2 involves an extra comparison.

2) For $1 < x < \tan \frac{11\pi}{24}$, Method No. 2 involves an extra addition and an extra division. As far as the number of comparisons are concerned, Method No. 2 has some advantage over Method No. 1.

3) Method No. 2 has an advantage in that sixteen instead of twenty-one stored constants are required.

5. ERROR PREDICTION FOR THE METHODS

In the following discussion of error prediction, use will be made of the following lemma.

Lemma: Let

$$N = (n_1\beta^{-1} + n_2\beta^{-2} + \ldots + n_\lambda\beta^{-\lambda} + \ldots)\beta^{\rho}$$

where $n_1 \neq 0$ be a given number,

and

$$\bar{N} = (n_1'\beta^{-1} + n_2'\beta^{-2} + \ldots + n_\lambda'\beta^{-\lambda})\beta^{\sigma}$$

where $n_1' \neq 0$ be an approximation for N. If

$$\left|\frac{\bar{N} - N}{N}\right| < a\,\beta^{-\tau} \qquad \tau \geq 1 \quad ,$$

then $\bar{N}$ differs from N by at most a units in the τ-th significant digit.

Proof: Since

$$\beta^{\rho-1} \leq N \leq \beta^{\rho} \quad ,$$

it follows that

$$|\bar{N} - N| < a\,\beta^{\rho-\tau} = (0\beta^{-1} + \ldots + 0\beta^{-(\tau-1)} + a\beta^{-\tau} + \ldots)\beta^{\rho}$$

Then

$$N - (0\beta^{-1} + \ldots + 0\beta^{-(\tau-1)} + a\beta^{-\tau} + \ldots)\beta^{\rho} < \bar{N} < N + (0\beta^{-1} + \ldots + 0\beta^{-(\tau-1)} + a\beta^{-\tau} + \ldots)\beta^{\rho}.$$

Hence, $\bar{N}$ differs from N by a units in the τ-th significant digit.

If $\tau > \lambda$, then, of course, all significant digits of $\bar{N}$ are correct.

For each method the error predictions will be based on using both nine and eight terms of the approximating polynomial. When nine terms are used, the truncation error is

$$\epsilon_T < (6)(10^{-22})\, x \quad \text{or} \quad (2^{-70})\, x \quad .$$

When eight terms are used, the truncation error is

$$\epsilon_{T'} < (4.5)(10^{-16})\, x \quad \text{or} \quad (2^{-51})\, x \quad .$$

Tables 2 and 3 show the error predictions for Method No. 1 using $\beta = 10$ and λ (the number of significant figures) ranging from 13 to 21. Table 2 is based on utilizing all nine terms of the approximating polynomial, whereas Table 3 is based on using only eight terms. The notation $\pm$ 2(12) means a possible error of $\pm$ 2 in the twelfth significant digit.

TABLE 2

(Method No. 1, $\beta = 10$, Nine terms)

	λ=13	λ=14	λ=15	λ=16	λ=17	λ=18	λ=19	λ=20	λ=21
$0 \le x < \tan \frac{\pi}{24}$	$\pm$2(12)	$\pm$2(13)	$\pm$2(14)	$\pm$2(15)	$\pm$2(16)	$\pm$2(17)	$\pm$2(18)	$\pm$2(19)	$\pm$2(20)
$\tan \frac{\pi}{24} \le x < \tan \frac{3\pi}{24}$	$\pm$7(12)	$\pm$7(13)	$\pm$7(14)	$\pm$7(15)	$\pm$7(16)	$\pm$7(17)	$\pm$7(18)	$\pm$7(19)	$\pm$7(20)
$\tan \frac{3\pi}{24} \le x < \tan \frac{5\pi}{24}$	$\pm$4(12)	$\pm$4(13)	$\pm$4(14)	$\pm$4(15)	$\pm$4(16)	$\pm$4(17)	$\pm$4(18)	$\pm$4(19)	$\pm$4(20)
$\tan \frac{5\pi}{24} \le x < \tan \frac{7\pi}{24}$	$\pm$3(12)	$\pm$3(13)	$\pm$3(14)	$\pm$3(15)	$\pm$3(16)	$\pm$3(17)	$\pm$3(18)	$\pm$3(19)	$\pm$3(20)
$\tan \frac{7\pi}{24} \le x < \tan \frac{9\pi}{24}$	$\pm$3(12)	$\pm$3(13)	$\pm$3(14)	$\pm$3(15)	$\pm$3(16)	$\pm$3(17)	$\pm$3(18)	$\pm$3(19)	$\pm$3(20)
$\tan \frac{9\pi}{24} \le x < \tan \frac{11\pi}{24}$	$\pm$3(12)	$\pm$3(13)	$\pm$3(14)	$\pm$3(15)	$\pm$3(16)	$\pm$3(17)	$\pm$3(18)	$\pm$3(19)	$\pm$3(20)
$\tan \frac{11\pi}{24} \le x < \infty$	$\pm$2(12)	$\pm$2(13)	$\pm$2(14)	$\pm$2(15)	$\pm$2(16)	$\pm$2(17)	$\pm$2(18)	$\pm$2(19)	$\pm$2(20)

TABLE 3

(Method No. 1, β = 10, Eight terms)

	λ=13	λ=14	λ=15	λ=16	λ=17	λ=18	λ=19	λ=20	λ=21
$0 \leq x < \tan \frac{\pi}{24}$	$\pm$2(12)	$\pm$2(13)	$\pm$2(14)	$\pm$2(15)	$\pm$6(16)	$\pm$5(16)	$\pm$5(16)	$\pm$5(16)	$\pm$5(16)
$\tan \frac{\pi}{24} \leq x < \tan \frac{3\pi}{24}$	$\pm$7(12)	$\pm$7(13)	$\pm$7(14)	$\pm$8(15)	$\pm$2(15)	$\pm$6(16)	$\pm$5(16)	$\pm$5(16)	$\pm$5(16)
$\tan \frac{3\pi}{24} \leq x < \tan \frac{5\pi}{24}$	$\pm$4(12)	$\pm$4(13)	$\pm$4(14)	$\pm$4(15)	$\pm$6(16)	$\pm$2(16)	$\pm$2(16)	$\pm$2(16)	$\pm$2(16)
$\tan \frac{5\pi}{24} \leq x < \tan \frac{7\pi}{24}$	$\pm$3(12)	$\pm$3(13)	$\pm$3(14)	$\pm$4(15)	$\pm$8(16)	$\pm$5(16)	$\pm$5(16)	$\pm$5(16)	+5(16)
$\tan \frac{7\pi}{24} \leq x < \tan \frac{9\pi}{24}$	$\pm$3(12)	$\pm$3(13)	$\pm$3(14)	$\pm$3(15)	$\pm$4(16)	$\pm$1(16)	$\pm$7(17)	$\pm$7(17)	$\pm$7(17)
$\tan \frac{9\pi}{24} \leq x < \tan \frac{11\pi}{24}$	$\pm$3(12)	$\pm$3(13)	$\pm$3(14)	$\pm$3(15)	$\pm$3(16)	$\pm$8(17)	$\pm$6(17)	$\pm$5(17)	$\pm$5(17)
$\tan \frac{11\pi}{24} \leq x < \infty$	$\pm$2(12)	$\pm$2(13)	$\pm$2(14)	$\pm$2(15)	$\pm$2(16)	$\pm$6(17)	$\pm$5(17)	$\pm$5(17)	$\pm$5(17)

TABLE 4

(Method No. 1, $\beta = 2$, Nine terms)

$/\lambda=$	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67
$0 \leq x < \tan\frac{\pi}{24}$	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64
$\tan\frac{\pi}{24} \leq x < \tan\frac{3\pi}{24}$	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62
$\tan\frac{3\pi}{24} \leq x < \tan\frac{5\pi}{24}$	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63
$\tan\frac{5\pi}{24} \leq x < \tan\frac{7\pi}{24}$	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63
$\tan\frac{7\pi}{24} \leq x < \tan\frac{9\pi}{24}$	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63
$\tan\frac{9\pi}{24} \leq x < \tan\frac{11\pi}{24}$	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63
$\tan\frac{11\pi}{24} \leq x < \infty$	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64

Tables 4 and 5 are again for Method No. 1 using $\beta = 2$ with λ ranging from 50 to 67. In these tables the entry shows the number of significant digits which are correct.

TABLE 5

(Method No. 1, $\beta = 2$, Eight terms)

$\lambda =$	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67
$0 \leq x < \tan \frac{\pi}{24}$	47	47	48	49	49	49	49	50	50	50	50	50	50	50	50	50	50	50
$\tan \frac{\pi}{24} \leq x < \tan \frac{3\pi}{24}$	45	46	47	47	48	49	49	49	49	50	50	50	50	50	50	50	50	50
$\tan \frac{3\pi}{24} \leq x < \tan \frac{5\pi}{24}$	46	47	48	49	49	50	50	51	51	51	51	51	51	51	51	51	51	51
$\tan \frac{5\pi}{24} \leq x < \tan \frac{7\pi}{24}$	46	47	48	49	50	50	51	51	52	52	52	52	52	52	52	52	52	52
$\tan \frac{7\pi}{24} \leq x < \tan \frac{9\pi}{24}$	46	47	48	49	50	51	51	52	52	52	52	52	52	52	52	52	52	52
$\tan \frac{9\pi}{24} \leq x < \tan \frac{11\pi}{24}$	46	47	48	49	50	51	52	52	52	53	53	53	53	53	53	53	53	53
$\tan \frac{11\pi}{24} \leq x < \infty$	47	48	49	50	51	52	52	53	53	53	53	53	53	53	53	53	53	53

Tables 6 through 9 show the error predictions for Method No. 2.

TABLE 6

(Method No. 2, β = 10, Nine terms)

	λ=13	λ=14	λ=15	λ=16	λ=17	λ=18	λ=19	λ=20	λ=21
$1 < x < \tan\frac{7\pi}{24}$	$\pm$5(12)	$\pm$5(13)	$\pm$5(14)	$\pm$5(15)	$\pm$5(16)	$\pm$5(17)	$\pm$5(18)	$\pm$5(19)	$\pm$5(20)
$\tan\frac{7\pi}{24} \leq x < \tan\frac{9\pi}{24}$	$\pm$5(12)	$\pm$5(13)	$\pm$5(14)	$\pm$5(15)	$\pm$5(16)	$\pm$5(17)	$\pm$5(18)	$\pm$5(19)	$\pm$5(20)
$\tan\frac{9\pi}{24} \leq x < \tan\frac{11\pi}{24}$	$\pm$4(12)	$\pm$4(13)	$\pm$4(14)	$\pm$4(15)	$\pm$4(16)	$\pm$4(17)	$\pm$4(18)	$\pm$4(19)	$\pm$4(20)

TABLE 7

(Method No. 2, β =10, Eight terms)

	λ=13	λ=14	λ=15	λ=16	λ=17	λ=18	λ=19	λ=20	λ=21
$1 < x < \tan\frac{7\pi}{24}$	$\pm$5(12)	$\pm$5(13)	$\pm$5(14)	$\pm$6(15)	$\pm$1(15)	$\pm$5(16)	$\pm$5(16)	$\pm$5(16)	$\pm$5(16)
$\tan\frac{7\pi}{24} \leq x < \tan\frac{9\pi}{24}$	$\pm$5(12)	$\pm$5(13)	$\pm$5(14)	$\pm$5(15)	$\pm$6(16)	$\pm$2(16)	$\pm$2(16)	$\pm$1(16)	$\pm$1(16)
$\tan\frac{9\pi}{24} \leq x < \tan\frac{11\pi}{24}$	$\pm$4(12)	$\pm$4(13)	$\pm$4(14)	$\pm$4(15)	$\pm$6(16)	$\pm$2(16)	$\pm$2(16)	$\pm$2(16)	$\pm$2(16)

TABLE 8

(Method No. 2, $\beta = 2$, Nine terms)

$/\lambda =$	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67
$1 < x < \tan \frac{7\pi}{24}$	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62
$\tan \frac{7\pi}{24} \leq x < \tan \frac{9\pi}{24}$	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63
$\tan \frac{9\pi}{24} \leq x < \tan \frac{11\pi}{24}$	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63

TABLE 9

(Method No. 2, $\beta = 2$, Eight terms)

$/\lambda =$	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67
$1 < x < \tan \frac{7\pi}{24}$	45	46	47	48	49	50	51	51	51	52	52	52	52	52	52	52	52	52
$\tan \frac{7\pi}{24} \leq x < \tan \frac{9\pi}{24}$	46	47	47	48	49	50	51	51	51	52	52	52	52	52	52	52	52	52
$\tan \frac{9\pi}{24} \leq x < \tan \frac{11\pi}{24}$	46	47	48	49	49	50	50	51	51	51	51	51	51	51	51	51	51	51

The following observations can be drawn from Tables 2 through 9.

1) In Method No. 1, if eight terms are used, then λ should be taken to be at most 16 when $\beta = 10$, and at most 57 when $\beta = 2$.

2) In Method No. 2, if eight terms are used, then λ should be at most 15 when $\beta = 10$, and at most 56 when $\beta = 2$.

3) The predicted errors for Method No. 2 are slightly larger in all cases than those for Method No. 1.

6. CONCLUSIONS

It seems that Method No. 1 is best both as to the number of operations performed and the error predictions. The only decided advantage of Method No. 2 is that fewer stored constants are required.

III. Study of Satellite Orbit Computation Methods and An Error Analysis of the Mathematical Procedures

Introduction

The objective of this study was to conduct a review of the various mathematical processes involved in numerical orbital computations with electronic computers; to ascertain the mathematical procedures best suited for orbit work to the degree of accuracy required; and to make an error analysis of these procedures. In addition, the purpose was to select from these methods the optimum procedures for speed and adaptability for use on the Remington Rand 1103-A Computer.

To initiate the study, preliminary discussions were arranged with recognized authorities in astronomy and others engaged in satellite orbit computation. Much of the existing literature was consulted, and in two instances the services of professional astronomers were engaged.

The first section of this paper contains outlines of the well-known methods of orbit determination, with attention directed to the mathematical procedures involved. The second section is a discussion of errors and a generalized procedure for analyzing these errors. In the third section an error analysis is presented for numerical integration procedures. The fourth section describes an orbit program which was prepared and used in assessing the errors in the method selected for orbit computation. The fifth section is devoted to comparisons and recommendations.

I. Methods of Orbit Computation.

The objective of the orbit computation procedure is to determine the location and velocity of the orbiting body as functions of time depending upon initial observations. Since it is not possible to obtain explicit expressions for these functions, suitable coordinates and velocity components of the orbiting body are given in the form of a tabular listing versus time (i.e., an ephemeris). This table can be constructed by integrating the equation of motion with the constants of integration determined initially and adjusted at intervals by means of observations.

This section contains an outline of the overall problem of satellite orbit computation.

A. The Procedure for Determining an Orbit.

Assuming that a sufficient number of observations are available, the procedure for computing an orbit may be divided into four main steps:

(a) obtaining the values of the parameters at some arbitrary time (the epoch);

(b) integration of the equations to obtain the computed positions;

(c) computation of the residuals, defined as the differences between the observed and computed values; and

(d) adjustment of the parameter values using the residuals.

For this study the method by which the parameters are obtained initially (e.g., from launch conditions) is not relevant; the values of these parameters are needed to initiate the integration process. The correction of the initial parameter values (or the correction of the values at some "anchor point" which have been derived from the initial values) is based on a differential corrections procedure, using the system of equations:

$$[O_{obs} - O_{comp}] = \sum_i \frac{\partial O_{comp}}{\partial P_i} \Delta P_i$$

where O_{obs} are the observations,

O_{comp} are the computed values of the quantities observed,

P_i are the appropriate parameters describing the motion (the constants of integration, orbital elements or functions of them) and

ΔP_i are the (unknown) corrections to the parameters.

The procedure (a) through (d) above is repeated until the residuals fall below an acceptable level. The process may fail if the starting parameters are greatly in error, or if gross observational errors are present, or if a poor choice of parameters has been made.

The differential correction process is to be the subject of a later report. This report is principally concerned with step (b) of the computing process, namely, the integration of the equations of motion.

B. The Equations of Motion.

The formulation of the equations of motion is outside the scope of this report. However, a discussion of the form of these equations and of their solutions is necessary to the remainder of this section and to later sections.

The classic methods of orbit computation, as applied to heliocentric orbits, are based upon the fact that the motion of celestial bodies is nearly the same as the motion in an idealized, two-body system. The motion of either body in the two-body problem is described by the differential equations (Ref 1. p. 144):

$$\frac{d^2x}{dt^2} = \frac{-k^2Mx}{r^3}$$

$$\frac{d^2y}{dt^2} = \frac{-k^2My}{r^3} \qquad (1\text{-}1)$$

$$\frac{d^2z}{dt^2} = \frac{-k^2Mz}{r^3}$$

where x, y, z are the coordinates of one body in a rectangular coordinate system with the origin at the center of the other body,

k^2 is the universal constant of gravitation,

M is the sum of the masses of the two bodies,[*] and

r is the distance between the centers of the two bodies.

The solution of equations (1-1) is a conic. For a given choice of initial conditions, x_o, y_o, z_o and $\dot{x}_o$, $\dot{y}_o$, $\dot{z}_o$ at $t = t_o$, the solution will be an ellipse with one focus (the perifocus) at the origin.

If any of the following conditions apply to the two-body problem:

(a) either of the two bodies is not perfectly spherical, or

(b) the mass of either body is not distributed in uniform, concentric shells about the center, or

(c) other bodies attract the two bodies, or

(d) either body moves in a resisting atmosphere, or

(e) any force acts on either body other than their mutual attractions,

the motion ceases to be a conic.

The solution for the two-body problem is useful, however, if the forces other than those present in the two-body problem are small. Modifying equations (1-1) to account for these disturbing forces (choosing units of length and time so that $k^2M = 1$), the system

$$\begin{aligned} \frac{d^2x}{d\tau^2} &= \frac{-x}{r^3} + F_x \\ \frac{d^2y}{d\tau^2} &= \frac{-y}{r^3} + F_y \qquad (1\text{-}2) \\ \frac{d^2y}{d\tau^2} &= \frac{-z}{r^3} + F_z \end{aligned}$$

[*] Because of uncertainties in the knowledge of the units of mass and length in the solar system in terms of accepted laboratory units, it is advantageous to select units of length and time such that k^2M is unity. For a discussion of this, see Ref. 2, p. 199 ff.

is obtained, where F_x, F_y, F_z are the disturbing forces acting parallel to the x, y, z axes. F_x, F_y, F_z are, in general, functions of the six coordinates and the time; the specification of F_x, F_y, F_z for earth satellites is a subject for research. In this report, only the components of F_x, F_y, F_z arising from (a) the non-spherical shape of the earth, and (b) atmospheric drag, will be considered.

Walters and Herrick (Ref. 3) give an expression for the potential of the earth, from which the following force components have been derived:

$$F_{x_{obs}} = -\frac{Jx}{r^5}\left(1 - 5\frac{z^2}{r^2}\right) - \frac{Hxz}{r^7}\left(3 - 7\frac{z^2}{r^2}\right) - \frac{Kx}{6r^7}\left(3 - 42\frac{z^2}{r^2} + 63\frac{z^4}{r^4}\right)$$

$$F_{y_{obs}} = F_{x_{ob}}\left(\frac{y}{x}\right) \qquad (1\text{-}3)$$

$$F_{z_{obs}} = -\frac{Jz}{r^5}\left(3 - 5\frac{z^2}{r^2}\right) + \frac{3H}{5r^5}\left(1 - 10\frac{z^2}{r^2} + \frac{35z^4}{3r^4}\right) - \frac{Kz}{6r^7}\left(15 - 70\frac{z^2}{r^2} + 63\frac{z^4}{r^4}\right)$$

where $F_{x_{ob}}$, $F_{y_{ob}}$, $F_{z_{ob}}$ are the components of the disturbing forces arising from the oblateness of the earth, and

J, H, K are the second, third, and fourth coefficients in the spherical harmonic expansion of the gravitational potential, with magnitudes:

$$J \sim 1632 \times 10^{-6}$$

$$H \sim 6 \times 10^{-6}$$

$$K \sim 9 \times 10^{-6}$$

x, y, z are a set of rectangular axes, with x, y in the equatorial plane, x pointing to the equinox, y pointing 90° east of x, and z coinciding with the polar axis, positive north.

Baker (Ref. 4)* gives the following expressions for the drag force, taking into account a rotating atmosphere:

*References in this section are to listings on pp.86-87.

$$
\begin{aligned}
F_{x_d} &= - \rho \frac{SC_d}{2m} V(\dot{x} + \omega_e y) \\
F_{y_d} &= - \rho \frac{SC_d}{2m} V(\dot{y} - \omega_e x) \qquad (1\text{-}4) \\
F_{z_d} &= - \rho \frac{SC_d}{2m} V \dot{z}
\end{aligned}
$$

where ρ is the density of the atmosphere,

S is the frontal area of the satellite,

C_d is the drag coefficient,

ω_e is the rotational velocity (sidereal) of the earth,

V is the magnitude of the velocity,

$V = \sqrt{\dot{x}^2 + \dot{y}^2 + \dot{z}^2}$, and

m is the mass of the vehicle.

It has been found that, for earth satellites with weight to frontal area ratios of the order of 5 - 10 (in English units) at altitudes in excess of 70 - 80 miles above the earth's surface, the drag and oblateness forces are less than the two-body forces by factors of several hundred. Under these conditions, the motion can be thought of as a conic (e.g., an ellipse) which is slowly changing form and orientation with time. The instantaneous values of x, y, z and $\dot{x}$, $\dot{y}$, $\dot{z}$ may be transformed into the elliptical elements (Appendix B) to obtain a succession of ellipses which contact the path; the total path is made up of a succession of infinitesimal arcs of these <u>osculating ellipses</u>.

To illustrate the concept of an ellipse with varying parameters, equations (1-2) were integrated numerically[*] for initial conditions corresponding to an earth satellite with an average altitude of 105 nautical miles. The density variation with altitude was assumed to be that given by the SAO Model Number Two (Ref. 5). The weight of the satellite was taken to be 2,000 pounds, the frontal area was assumed to be 30 square feet,

[*] Using the test program described in Section 4 .

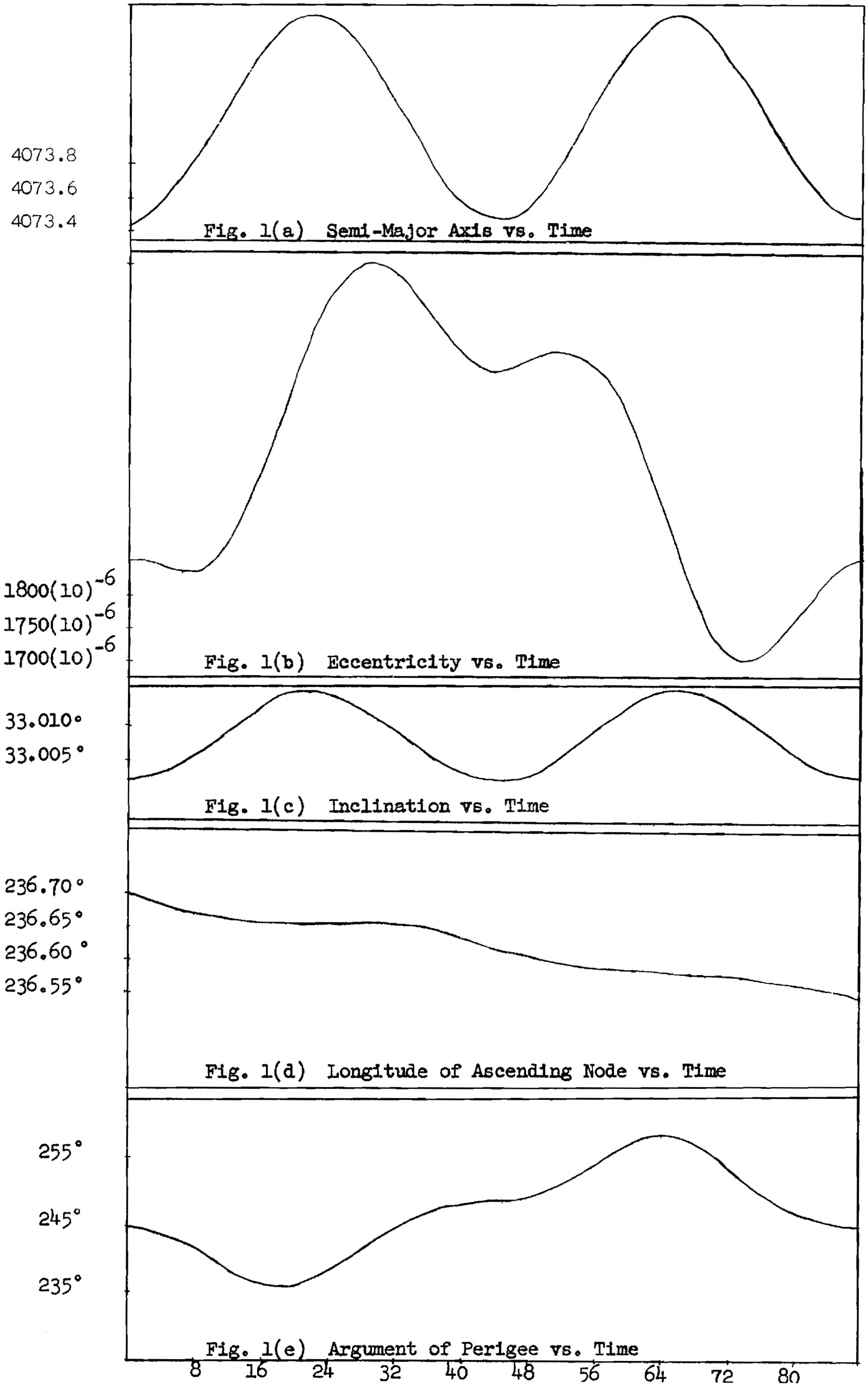

Fig. 1(a) Semi-Major Axis vs. Time

Fig. 1(b) Eccentricity vs. Time

Fig. 1(c) Inclination vs. Time

Fig. 1(d) Longitude of Ascending Node vs. Time

Fig. 1(e) Argument of Perigee vs. Time

and a constant drag coefficient of 2 was used. The values of x, y, z, $\dot{x}$, $\dot{y}$, $\dot{z}$ obtained from the integration were transformed into the standard elements of the ellipse (Appendix B). The magnitudes of five of the elements are shown plotted against time for one revolution of the satellite in Figure 1. Figure 1(a) shows a plot of the instantaneous values of the semi-major axis, in statute miles, as a function of time. Figure 1(b) shows a plot of the values of the eccentricity as a function of time. Figure 1(c) shows a plot of the values of the inclination in degrees as a function of time, and Figure 1(d) shows a plot of the right ascension of the ascending node in degrees as a function of time. The plot of the right ascension of the ascending node shows the short period terms superposed on the characteristic constant rate of change (precession) clearly. In Figure 1(e) the values of the argument of perigee in degrees as a function of time are shown. For the small eccentricities corresponding to the initial conditions selected, the argument of perigee varies rapidly with position in the orbit.

The concept of the conic with varying parameters leads naturally to the application of a mathematical procedure known as the Variation of Elements, or Variation of Parameters.

C. Variation of Parameters: Standard Elements.

The coordinate system is shown in Figure 2. The notation follows that used by Herrick (Ref. 5).

$\vec{I}$, $\vec{J}$, $\vec{K}$ form a right-handed system of unit vectors, coinciding with the x, y, z axes.

$\vec{P}$, $\vec{Q}$, $\vec{W}$ form a right-handed system of unit vectors, with $\vec{P}$ pointing to perigee, $\vec{Q}$ pointing 90° east of $\vec{P}$ in the plane of the orbit, $\vec{W}$ pointing in a direction normal to the orbit plane.

$\vec{U}$, $\vec{V}$, $\vec{W}$ form a right-handed system of unit vectors, with $\vec{U}$ pointing to the satellite, $\vec{V}$ pointing 90° east of $\vec{U}$ in the plane of the orbit.

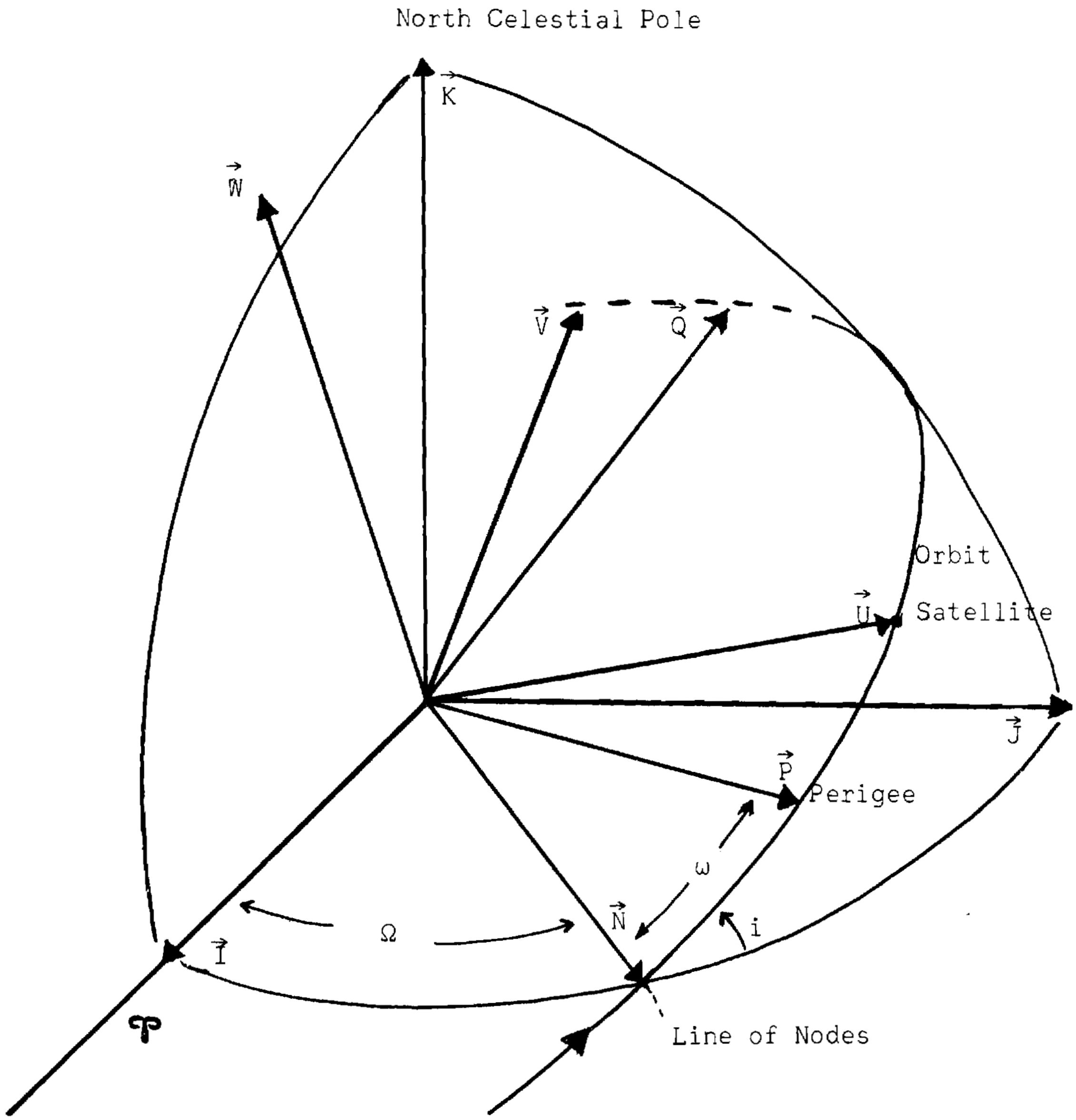
North Celestial Pole
$\vec{K}$
$\vec{W}$
$\vec{V}$
$\vec{Q}$
Orbit
$\vec{U}$
Satellite
$\vec{J}$
$\vec{P}$
Perigee
ω
Ω
$\vec{N}$
i
$\vec{I}$
♈
Line of Nodes

Figure 2.

The perturbing forces F_x, F_y, F_z are transformed into components parallel to the $\vec{P}$, $\vec{Q}$, $\vec{W}$, $\vec{U}$, $\vec{V}$ vectors by the equations:

$$\begin{aligned} F_P &= \vec{F} \cdot \vec{P} \\ F_Q &= \vec{F} \cdot \vec{Q} \\ F_W &= \vec{F} \cdot \vec{W} \\ F_U &= \vec{F} \cdot \vec{U} \\ F_V &= \vec{F} \cdot \vec{V} \end{aligned} \tag{1-5}$$

where

$$\begin{aligned} \vec{F} &= \vec{I}F_x + \vec{J}F_y + \vec{K}F_z \ , \\ \vec{P} &= \vec{I}P_x + \vec{J}P_y + \vec{K}P_z \ , \end{aligned} \tag{1-6}$$

etc. for $\vec{Q}$, $\vec{W}$, $\vec{U}$, $\vec{V}$.

The components of $\vec{P}$, $\vec{Q}$, $\vec{W}$, $\vec{U}$, $\vec{V}$ parallel to x, y, z axes are given by the set of equations:

$$\begin{aligned} \vec{P} &= \vec{U} \cos v - \vec{V} \sin v \\ \vec{Q} &= \vec{U} \sin v + \vec{V} \cos v \\ \vec{W} &= \vec{P} \times \vec{Q} \end{aligned}$$

where

$$\begin{aligned} \vec{U} &= \frac{\vec{r}}{r} \\ \sqrt{p}\,\vec{V} &= r\dot{\vec{r}} - \dot{r}\vec{r} \\ p &= (r\dot{\vec{r}} - \dot{r}\vec{r}) \cdot (r\dot{\vec{r}} - \dot{r}\vec{r}) \\ \vec{r} &= \vec{I}x + \vec{J}y + \vec{K}z \\ \dot{\vec{r}} &= \vec{I}\dot{x} + \vec{J}\dot{y} + \vec{K}\dot{z} \\ \dot{r} &= \frac{\vec{r} \cdot \dot{\vec{r}}}{r} \ , \qquad \dot{r} \neq |\dot{\vec{r}}| \\ e \sin v &= \dot{r}\sqrt{p} \\ e \cos v &= \frac{p}{r} - 1 \end{aligned} \tag{1-7}$$

In these equations,

$\sqrt{p}$ = the angular momentum about the geocenter,

e = the eccentricity,

v = the true anomaly,

and the other quantities are as previously defined.

The differential equations for the variation of the parameters are (Ref. 6):

$$
\begin{aligned}
\frac{di}{d\tau} &= F_W \frac{r \cos u}{a\sqrt{1 - e^2}} \\
\frac{d\Omega}{d\tau} &= F_W \frac{r \sin u}{a\sqrt{1 - e^2} \sin i} \\
\frac{d\omega}{d\tau} &= F_V \frac{r \sin v}{ae\sqrt{1 - e^2}} - F_P \frac{\sqrt{1 - e^2}}{e} - F_W \frac{r \sin u}{a\sqrt{1 - e^2}} \cot i \\
\frac{de}{d\tau} &= \frac{r}{a} F_V[e + \cos v] + F_Q \\
\frac{dn}{d\tau} &= \frac{-3n^{\circ} e}{\sqrt{1 - e^2}} F_Q - \frac{3n^{\circ}}{\sqrt{1 - e^2}} F_V \\
\frac{dM}{d\tau} &= \frac{-r}{ae} F_V \sin v + \frac{(1 - e^2)}{e} F_P - \frac{2r}{a} F_U + n
\end{aligned}
\qquad (1\text{-}8)
$$

where

u = argument of latitude, $u = v + \omega$,

n° = the mean angular motion for the unperturbed orbit, defined by

$$n^{\circ} = \frac{2\pi}{T} = a^{-3/2},$$

T is the anomalistic period in the time units chosen,

M = the mean anomaly.

The other quantities have been defined previously.

The equations (1-8) are to be integrated numerically, with the forces parallel to the $\vec{P}$, $\vec{Q}$, $\vec{W}$, $\vec{U}$, $\vec{V}$ directions determined at each step of the integration by equations (1-7). Equations (1-5) and (1-7) are well adapted to machine computation, since no trigonometric routines appear in the equations for transforming the forces.

The equations (1-8) will fail for small eccentricities, because of the appearance of the eccentricity in the denominator of the equations for the mean anomaly and the argument of perigee. The failure of the equation for the argument of perigee is to be expected from the choice of the parameters:

perigee location is indeterminate for circular orbits.

The appearance of the term, $\frac{F_W}{\sin i}$, in the equations for the right ascension of the ascending node and for the argument of perigee is another source of possible difficulty for near equatorial orbits. Asymmetry between the northern and southern hemispheres, as reflected in the presence of the odd harmonics in the spherical harmonic expansion of the gravitational potential, may cause this term to "blow up" for near equatorial orbits.

Other methods, which differ from the method of Variation of Parameters described above in the choice of parameters are:

(a) Stromgren's Method, in which the elements i, Ω, ω are replaced by components of the $\vec{W}$, $\vec{P}$ vectors (Ref. 6), with time as the independent variable.

(b) Oppolzer's Method, in which the eccentric anomaly is the independent variable, thus avoiding the solution of Kepler's equation at each step in the integration (Ref. 6).

(c) Merton's Method, in which the mean anomaly is the independent variable.

For the purposes of this report, these methods offer little or nothing over the standard method outlined in equations (1-8) ; each suffers from the same difficulties for quasi-circular and near equatorial orbits.

In Hansen's Method the position of the object is referred to the original osculating orbit plane. The radius vector and the true anomaly are perturbed. The distance of the object from the osculating plane, the change in the projection of the radius vector on the osculating plane, and the rotation of the line of apsides are obtained by integration. Cylindrical polar coordinates are used (Ref. 6, 7, 8). A modification of Hansen's Method has been applied by Herget and co-workers to the computation of the orbits of earth satellites under the Vanguard program, but a description of the method has not been published.

Herrick (Ref. 9) has proposed sets of parameters for earth satellites which avoid the difficulties associated with near-equatorial and quasi-circular orbits.

D. Herrick's Method: Low eccentricity orbits.

Herrick's Method differs from the methods above in that a special set of parameters is chosen. The application of Herrick's Method to heliocentric orbits is described in Ref. 6; the following has been proposed for small eccentricity orbits for near earth satellites (Ref. 9).

The components of the unit vectors $\vec{P}$, $\vec{Q}$, $\vec{U}$, $\vec{V}$, $\vec{W}$ are found as in equations (1-7). The semi-major axis is found from:

$$a = \frac{P}{1 - e^2}$$

$$e^2 = P\dot{r}^2 + \left(\frac{P}{r} - 1\right)^2 \tag{1-9}$$

The mean longitude of the object, L_o, is found from the following equations:

$$L_o = M_o + \pi$$

$$M_o = E - e \sin E$$

$$\sin E = \frac{\sqrt{1 - e^2} \sin v}{(1 + e \cos v)}$$

$$\cos E = \frac{\cos v + e}{1 + e \cos v} \tag{1-10}$$

$$\cos \pi = \frac{P_x + Q_y}{1 + W_z}$$

$$\sin \pi = \frac{P_y - Q_x}{1 + W_z}$$

The following intermediate quantities are formed from the disturbing force, F_x, F_y, F_z, where the grave (′) indicates a disturbed quantity:

$$
\begin{aligned}
D' &= xF_x + yF_y + zF_z \\
\dot{D}' &= 2(\dot{x}F_x + \dot{y}F_y + \dot{z}F_z) \\
a'_x &= (\dot{D}'x - D'\dot{x} - DF_x) \\
a'_y &= (\dot{D}'y - D'\dot{y} - DF_y) \\
a'_z &= (\dot{D}'z - D'\dot{z} - DF_z) \\
h'_x &= (yF_z - zF_y) \\
h'_y &= (zF_x - xF_z) \\
h'_z &= (xF_y - yF_x)
\end{aligned}
\tag{1-11}
$$

The perturbation in the mean longitude is formed from the following equations:

$$
\begin{aligned}
ev' &= -Q_x a'_x - Q_y a'_y - Q_z a'_z \\
rb' &= W_x F_x + W_y F_y + W_z F_z \\
\ell' &= \frac{z(r\dot{b}')}{(1 + W_z)\sqrt{P}} \\
L' &= \ell' - \frac{2D'}{\sqrt{a}} - \frac{e^2 v'}{1 + \sqrt{1 - e^2}} \\
n' &= -\frac{3}{2}\, n\, a\, \dot{D}'
\end{aligned}
\tag{1-12}
$$

In the following equations,

$$
\begin{aligned}
\vec{h} &= \vec{h}_o + \int \vec{h}'\, d\tau \\
\vec{a} &= \vec{a}_o + \int \vec{a}'\, d\tau \\
n &= n_o + \int n'\, d\tau \\
L &= L_o + n_o\, \Delta\tau + \int\int n'\, d\tau^2 + \int L'\, d\tau
\end{aligned}
\tag{1-13}
$$

the integrations are performed numerically, where

$\vec{h}$ is the vector component of the angular momentum,

$$\vec{h} = \sqrt{p}\,\vec{W} ,$$

$\vec{h}'$ is the disturbed angular momentum, whose components are given by equations (1-11),

$\vec{a}$ is the vector of magnitude e, which locates perigee, defined by

$$\vec{a} = e\vec{P} ,$$

$\vec{a}'$ is the disturbed component of this vector, whose components are given by equation (1-10),

L is the mean longitude of the object, defined by

$$L = M + \pi ,$$

π is defined by $\pi = \Omega + \omega$, and

D, b, and ℓ are to regarded as intermediate quantities (for the purposes of this description), defined in equations (1-11) and (1-12).

The equations within the integration loop contain a minimum number of trigonometric routines. The presence of the eccentricity in the denominator of the mean anomaly integral is avoided. However, these gains are made at the expense of complex equations and more computation within the integration loop.

E. Encke's Method.

This method differs from those described above in that the integration is performed in rectangular coordinates. The quantities integrated are the differences between the coordinates of the disturbed satellite and the coordinates that the satellite would have had in the absence of the disturbing forces. These differences are defined by

$$\begin{aligned} \xi &= x - x^{\circ} \\ \eta &= y - y^{\circ} \\ \zeta &= z - z^{\circ} \end{aligned} \tag{1-14}$$

where x°, y°, z° are the coordinates of the equivalent undisturbed motion (i.e., the two-body solution).

Substituting equations (1-14) into (1-2) and reducing:

$$\ddot{\xi} = \frac{1}{r^{o3}}\left[x\left(1 - \frac{r^{o3}}{r}\right) - \xi\right] + F_x$$

$$\ddot{\eta} = \frac{1}{r^{o3}}\left[y\left(1 - \frac{r^{o3}}{r}\right) - \eta\right] + F_y \qquad (1\text{-}15)$$

$$\ddot{\zeta} = \frac{1}{r^{o3}}\left[z\left(1 - \frac{r^{o3}}{r}\right) - \zeta\right] + F_z$$

Substituting (1-14) into the equations for the disturbed values of r,

$$r^2 = x^2 + y^2 + z^2$$

and expanding in a Taylor's series, yields

$$\left(\frac{r}{r^{o3}}\right) = (1 + 2q)$$

where q is the term in the expansion

$$q = \frac{1}{r^{o3}}\left[(x^o + \tfrac{1}{2}\xi)\xi + (y^o + \tfrac{1}{2}\eta)\eta + (z^o + \tfrac{1}{2}\zeta)\zeta\right]$$

Finally

$$1 - \left(\frac{r^o}{r}\right)^3 = 1 - (1 + 2q)^{-3/2} = 3q - \frac{3.5}{2\,!}q^2 + \ldots = fq$$

The equations (1-15) now take the form:

$$\ddot{\xi} = \frac{1}{r^{o3}}\left[fqx - \xi\right] + F_x$$

$$\ddot{\eta} = \frac{1}{r^{o3}}\left[fqy - \eta\right] + F_y$$

$$\ddot{\zeta} = \frac{1}{r^{o3}}\left[fqz - \zeta\right] + F_z$$

The error in the accelerations, when the perturbing forces change slowly with position, is approximately equal to the errors in the series expansions for fq divided by r^{o3}. When the departure from the two-body path exceeds certain limits, this error will be excessive; in practice, a new reference orbit is calculated before this happens.

F. Cowell's Method*

In this method the fundamental equations of motion are integrated directly by means of the formulae:

$$\delta^{-1} X_{\frac{1}{2}} = h\dot{x}_o + \frac{1}{2} X_o + \frac{1}{12} \mu\delta X_o - \frac{11}{720} \mu\delta^3 X_o + \ldots,$$

$$\delta^{-2} X_o = x_o - \frac{1}{12} X_o + \frac{1}{240} \delta^2 X_o - \frac{31}{60480} \delta^4 X_o + \ldots, \quad \text{and}$$

$$x_o = \delta^{-2} X_o + \frac{1}{12} X_o - \frac{1}{240} \delta^2 X_o + \frac{31}{60480} \delta^4 X_o + \ldots,$$

where

$$X_p = h^2 \ddot{x}_p \quad ,$$

and x_o and $\dot{x}_o$ are the values of the double and single integrations of X respectively, and h is the time step. The symbols δ^i, $i = \pm 1, \pm 2, \ldots$ and μ are the central difference and averages operators from finite difference theory (Ref. 10, 11, 12). There are similar equations for Y and Z.

Cowell's Method does not use the solution of the two-body problem in the integration of equations (1-2). Because the entire equation is integrated directly, a small time interval must be used per step to ensure a small truncation error. In addition, the use of central differences requires that a table of higher differences be constructed as the integration proceeds. The construction of this table from the initial conditions requires some type of iterative scheme, in other words, the procedure is not "self-starting." Because the procedure is not self-starting, it is not as convenient to change integration step time intervals in Cowell's Method as in certain other numerical integration procedures (e.g., Runge-Kutta). Despite these apparent disadvantages, however, the formulae are simple, the use of

---

* The method of integrating directly for the rectangular coordinates is usually referred to as "Cowell's Method" (for example, see Ref. 13). However, as pointed out in Ref. 8, p. 98, and in Ref. 14, in the strict sense Cowell's Method is an integration of the equations of motion by a method of central differences. In this report, the term "Cowell's Method" is used to mean the direct integration of equations (1-2) by central differences of any order.

rectangular coordinates offers advantages in the construction of an ephemeris and in the computation of residuals, and the problem of starting the integration can be overcome by using one of the self-starting numerical integration procedures in combination with an iterative scheme.

G. Comments on the Selection of a Method for Orbit Determination for Earth Satellites.

In the time of the hand computation of heliocentric orbits, great emphasis was placed on methods which combined a minimum of computing labor with all possible long term prediction accuracy. Because, in general, the perturbing forces in heliocentric orbits are extremely small, methods utilizing the two-body solution minimize computational labor and consequently were favored.

The problem of determining the orbits of near-earth satellites, and the advent of high-speed electronic computers make it desirable to re-evaluate the concept of orbit determination. The perturbing forces on near-earth satellites are much greater, generally, than the perturbing forces encountered in heliocentric orbits. The time scale is much smaller in terms of the lifetime of the satellites and in the need for immediate publication of information. The premium on minimized computation is no longer as great since the computing machine has stamina and speed. However, in the era of hand computing, all the steps in the computing procedure received the personal attention and judgment of the analyst, and procedures which exhibited instabilities or other danger signals were modified or replaced as the computation progressed. The success of the procedure and the validity of the result frequently reflected the agility of the person doing the computation. High-speed computing machinery has little choice and virtually no judgment. Because of this, and because of problems in programming and program checking, modern practice tends toward the use of a few well-tested programs for a large number of situations. In many cases, speed may be sacrificed to some extent for dependability of results.

The optimum orbit determination procedure, therefore, should be applicable to a wide variety of situations, have good overall running time features, and be capable of yielding a high quality result.

2. Discussion of Errors

A mathematical procedure common to all the methods of Section I is numerical integration. Hence, it will be necessary to study and to evaluate the errors involved in such a procedure. Regardless of the specific integration technique, the procedure is in essence one of approximating quantities given by infinite series by truncated (finite) series. Therefore, in every operation in which this approximation is made, a truncation error will be introduced. In machine computation it is not feasible to try to correct this error at each step in the procedure, and if it is not controlled this error may grow over long-time periods to the point of ultimately casting doubt on the validity of any digit retained in the solution.

Another type of error (round-off) exists and depends on the control of truncation error. The computing machine can carry only a certain number of digits and every entry used must be rounded off to this number of digits. This is unavoidable, and the only recourse is to attempt to evaluate the rate of growth of these errors and plan the machine procedures in such a way that results are of the desired accuracy.

It can be stated generally that the above errors behave in an antagonistic manner. Any action taken to reduce the truncation error tends to increase the round-off error and vice versa.

A third type of error (propagated error) due to inaccuracies in the measurements determining the initial conditions can be handled in a manner similar to the method for treating the truncation error. The difference in the two methods is simply that in treating truncation error the initial values for the solution of the equations of variation are taken to be zero, where in treating propagated error the initial values for the solution of the equations of variation are taken to be the estimated errors in the initial coordinates. A general procedure for estimating errors is based on the solutions of the system of equations adjoint to the variational equation of the particular problem. In outline the method follows.

Let

$$y_i \;=\; f_i(y_1, y_2, \ldots, y_n) \;, \qquad (i = 1,2, \ldots, n) \qquad (2\text{-}1)$$

be a given set of n first order differential equations. A solution is sought subject to the initial conditions:

$$y_i(0) \;=\; y_{io} \;, \qquad (i = 1,2, \ldots, n) \qquad (2\text{-}2)$$

Let the solution corresponding to these initial conditions be denoted by y_i , $(i = 1,2, \ldots, n)$. If ξ_i represents a neighboring solution of the system of differential equations such that the differences $\xi_i - y_i$, $(i = 1,2, \ldots, n)$ are sufficiently small so that their squares and higher powers may be neglected, then these differences satisfy a system of differential equations (the equations of variations) of the form

$$\frac{d}{dt}(\xi_i - y_i) \;=\; \sum_{j=1}^{n} a_{ij}\,(\xi_j - y_j) \qquad (2\text{-}3)$$

The right hand members of the preceding system of differential equations are obtained by expanding the functions $f_i(y_i, y_2, \ldots, y_n)$ into Taylor series through the first order terms.

Suppose now that some numerical step-by-step method of integration were used to solve the system (2-1) subject to the initial conditions (2-2). Let η_i denote this solution. Now η_i do not satisfy the system (2-1), but a system which can be written as

$$\eta_i \;=\; f_i(\eta_1, \eta_2, \ldots, \eta_n) + b_i(t) \;, \qquad (i = 1,2, \ldots, n) \qquad (2\text{-}4)$$

The differences $\eta_i - y_i$ will then satisfy the equations of variation

$$\frac{d}{dt}(\eta_i - y_i) \;=\; \sum_{j=1}^{n} a_{ij}\,(\eta_j - y_j) + b_i \qquad (2\text{-}5)$$

The initial values of $\eta_i - y_i$ are zero, but at the end of one step in the numerical integration procedure the final values will not be zero due to the fact that the $b_i(t)$ are not zero in the interval. The $b_i(t)$ may be different from zero, because of truncation error, round-off error, or both. The method to be described expresses the final values of the differences in $\eta_i - y_i$ in terms of the $b_i(t)$ and certain solutions of the system of differential equations adjoint to the equations of variations (2-3) (Ref. 17).

Let the system (2-5) be:

$$\dot{x}_i = \sum_{j=1}^{n} a_{ij}(t)\, x_j + b_i(t) \quad , \qquad (i = 1, 2, \ldots, n)$$

where the $a_{ij}(t)$ and $b_i(t)$ are functions of time. Let the system of differential equations adjoint to (2-3) be:

$$-\dot{\lambda}_i = \sum_{j=1}^{n} a_{ji}(t)\, \lambda_i \quad . \qquad (2\text{-}6)$$

Now

$$\frac{d}{dt}\Big(\sum_{j=1}^{n} x_j\, \lambda_j\Big) = \sum_{j=1}^{n} \dot{x}_j\, \lambda_j + \sum_{j=1}^{n} x_j\, \dot{\lambda}_j \quad ,$$

$$= \sum_{j=1}^{n} \Big(\sum_{k=1}^{n} a_{jk}(t) \quad x_k + b_j\Big) \lambda_j - \sum_{j=1}^{n} \Big(\sum_{k=1}^{n} a_{kj}(t)\, \lambda_k\Big) x_j \quad ,$$

$$= \sum_{k=1}^{n} b_k(t)\, \lambda_k \quad .$$

Upon integration between the limits t to $t + h$, the equation

$$\sum_{i=1}^{n} x_i(t + h)\, \lambda_i(t + h) - \sum_{i=1}^{n} x_i(t)\, \lambda_i(t) = \int_t^{t+h} \sum_{i=1}^{n} b_i(t)\, \lambda_i(t)\, dt$$

is obtained.

If the solution of the system (2-6) is chosen so that $\lambda_i(t+h) = 1$, $\lambda_j(t+h) = 0$, $(j \neq i)$, then

$$x_i(t + h) = \int_t^{t+h} \sum_{k=1}^{n} b_k(t)\, \lambda_k\, dt \; .$$

This equation expresses the error in the variable y_i for a single time step of length h in terms of the quantities $b_i(t)$ and a solution of the adjoint system (2-6) If several time steps are taken, then

$$x_i(t + ph) = \int_t^{t+ph} \sum_{k=1}^{n} b_k(t) \lambda_k \, dt$$

where again $\lambda_i(t+ph) = 1$, $\lambda_j(t+ph) = 0$, $(j \neq 1)$.

In applying these results to a system of differential equations, the $b_i(t)$ are determined by the truncation error, round-off error, or both. The solution of the adjoint system (2-6) need not be highly accurate for an estimate of the errors in the original variables.

As indicated on page 49, the propagated error is determined in the same manner as above with $b_i(t) = 0$ and the values of $x_i(t)$ taken as the estimates of the errors in the initial values of the coordinates. Thus the propagated error in the coordinate x_i is given by:

$$x_i(t + ph) = \sum_{j=1}^{n} x_j(t) \lambda_j(t)$$

where again the solution of the adjoint system (2-6) is determined by the initial conditions $\lambda_i(t + ph) = 1$, $\lambda_j(t + ph) = 0$, $(j \neq i)$, and the $x_j(t)$ are the estimates of the errors in the initial values of the coordinates.

This procedure will now be applied to the equations of motion of an earth satellite.

3 Error Analysis for Numerical Integration Method.

A. One form of the equations of motion for an earth satellite (Ref. 16, p. 12-15) is:

$$\begin{aligned}
\ddot{x} &= -\frac{k^2 M_\oplus}{r^2}\left\{1 - 5JR_e^2\left(\frac{z^2}{r^2} - \frac{1}{5}\right)\right\}\frac{x}{r} + C_d\sigma\,\theta\,V\,\dot{x} \\
\ddot{y} &= -\frac{k^2 M_\oplus}{r^2}\left\{1 - 5JR_e^2\left(\frac{z^2}{r^2} - \frac{1}{5}\right)\right\}\frac{y}{r} + C_d\sigma\,\theta\,V\,\dot{y} \qquad (3\text{-}1) \\
\ddot{z} &= -\frac{k^2 M_\oplus}{r^2}\left\{1 - 5JR_e^2\left(\frac{z^2}{r^2} - \frac{3}{5}\right)\right\}\frac{z}{r} + C_d\sigma\,\theta\,V\,\dot{z}
\end{aligned}$$

where x, y, z are rectangular coordinates with origin at center of the earth and the xy-plane is that of the equator with x-axis directed toward the vernal equinox; $r^2 = x^2 + y^2 + z^2$; $V^2 = \dot{x}^2 + \dot{y}^2 + \dot{z}^2$; $k^2 = 3.71941900 \times 10^9$; $M_\oplus$ is the mass of the earth; J is an oblateness constant; R_e is the equatorial radius of the earth; C_d is a drag coefficient; σ is atmospheric density; θ is a conversion factor. Substitution of $a = -k^2 M_\oplus$, $b = -5JR_e^2$, $c = -\frac{1}{5}$, $d = -\frac{3}{5}$, and $H = C_d\sigma\,\theta$ reduces this system to:

$$\begin{aligned}
\ddot{x} &= \frac{ax}{r^3} + \frac{abcx}{r^5} + \frac{abxz^2}{r^7} + HV\dot{x} \\
\ddot{y} &= \frac{ay}{r^3} + \frac{abcy}{r^5} + \frac{abyz^2}{r^7} + HV\dot{y} \qquad (3\text{-}2) \\
\ddot{z} &= \frac{az}{r^3} + \frac{abdz}{r^5} + \frac{abz^3}{r^7} + HV\dot{z}\ .
\end{aligned}$$

Introducing new variables u, v, and w, the system (3-2) reduces to the following system of six first order equations:

$$
\begin{aligned}
\dot{x} &= u \\
\dot{y} &= v \\
\dot{z} &= w \\
\dot{u} &= \frac{ax}{r^3} + \frac{abcx}{r^5} + \frac{abxz^2}{r^7} + H\,V\,\dot{x} \\
\dot{v} &= \frac{ay}{r^3} + \frac{abcy}{r^5} + \underline{abyz^2} + H\,V\,\dot{y} \\
\dot{w} &= \frac{az}{r^3} + \frac{abdz}{r^5} + \frac{abz^3}{r^7} + H\,V\,\dot{z}
\end{aligned}
\qquad (3\text{-}3)
$$

A solution for $x(t)$, $y(t)$, $z(t)$, $u(t)$, $v(t)$, $w(t)$ is sought subject to the initial conditions $x_o = x(t_o)$, $y_o = y(t_o)$, $z_o = z(t_o)$, $u_o = u(t_o)$, $v_o = v(t_o)$, and $w_o = w(t_o)$ at time $t = t_o$. Suppose that $\xi(t)$, $\eta(t)$, $\alpha(t)$, $\beta(t)$, $\gamma(t)$, $\delta(t)$ is a solution other than $x(t)$, $y(t)$, $z(t)$, $u(t)$, $v(t)$, and $w(t)$, such that $|\xi(t) - x(t)|$, $|\eta(t) - y(t)|$, $|\alpha(t) - z(t)|$, $|\beta(t) - u(t)|$, $|\gamma(t) - v(t)|$, and $|\delta(t) - w(t)|$ are so small that squares and higher order terms can be neglected. Using Taylor's Theorem and neglecting higher order terms in $(\xi - x)$, $(\eta - y)$, $(\alpha - z)$, $(\beta - u)$, $(\gamma - v)$, and $(\delta - w)$ the following are derived as the equations of variation (Ref. 11):

$$
\begin{aligned}
\frac{d}{dt}(\xi - x) &= \beta - u \\
\frac{d}{dt}(\eta - y) &= \gamma - v \\
\frac{d}{dt}(\alpha - z) &= \delta - w
\end{aligned}
$$

$$
\begin{aligned}
\frac{d}{dt}(\beta - u) = {} & (\xi - x)\left\{\frac{a(r^2 - 3x^2)}{r^5} + \frac{abc(r^2 - 5x^2)}{r^7} + \frac{abz^2(r^2 - 7x^2)}{r^9}\right\} \\
& + (\eta - y)\left\{\frac{a(-3xy)}{r^5} + \frac{abc(-5xy)}{r^7} + \frac{abz^2(-7xy)}{r^9}\right\} +
\end{aligned}
$$

$$+ \quad (\alpha - z) \left\{ \frac{a(- 3xz)}{r^5} + \frac{abc(- 5xz)}{r^7} + \frac{abz(2r^2 - 7z^2)}{r^9} \right\}$$

$$+ \quad (\beta - u) \left\{ \frac{H(2u^2 + v^2 + w^2)}{(u^2 + v^2 + w^2)^{\frac{1}{2}}} \right\} + (\gamma - v) \left\{ \frac{H\ u\ v}{(u^2 + v^2 + w^2)^{\frac{1}{2}}} \right\} + (\delta - w) \left\{ \frac{H\ v\ w}{(u^2 + v^2 + w^2)^{\frac{1}{2}}} \right\} ,$$

$$\frac{d}{dt} (\gamma - v) = (\xi - x) \left\{ \frac{a(- 3xy)}{r^5} + \frac{abc(- 5xy)}{r^7} + \frac{abz^2(- 7xy)}{r^9} \right\} \tag{3-4}$$

$$+ \quad (\eta - y) \left\{ \frac{a(r^2 - 3y^2)}{r^5} + \frac{abc(r^2 - 5y^2)}{r^7} + \frac{abz^2(r^2 - 7y^2)}{r^9} \right\}$$

$$+ \quad (\alpha - z) \left\{ \frac{a(- 3xy)}{r^5} + \frac{abc(- 5yz)}{r^7} + \frac{abyz(2r^2 - 7z^2)}{r^9} \right\}$$

$$+ \quad (\beta - u) \left\{ \frac{H\ u\ v}{(u^2 + v^2 + w^2)^{\frac{1}{2}}} \right\} + (\gamma - v) \left\{ \frac{H(u^2 + 2v^2 + w^2)}{(u^2 + v^2 + w^2)^{\frac{1}{2}}} \right\} + (\delta - w) \left\{ \frac{H\ v\ w}{(u^2 + v^2 + w^2)^{\frac{1}{2}}} \right\} ,$$

$$\frac{d}{dt} (\delta - w) = (\xi - x) \left\{ \frac{a(- 3xz)}{r^5} + \frac{abd(- 5xz)}{r^7} + \frac{abz^2(- 7xz)}{r^9} \right\}$$

$$+ \quad (\eta - y) \left\{ \frac{a(- 3yz)}{r^5} + \frac{abd(- 5yz)}{r^7} + \frac{abz^2(- 7yz)}{r^9} \right\}$$

$$+ \quad (\alpha - z) \left\{ \frac{a(r^2 - 3z^2)}{r^5} + \frac{abd(r^2 - 5z^2)}{r^7} + \frac{abz^2(3r^2 - 7z^2)}{r^9} \right\}$$

$$+ \quad (B - u) \left\{ \frac{H\ u\ v}{(u^2 + v^2 + w^2)^{\frac{1}{2}}} \right\} + (\gamma - v) \left\{ \frac{H\ v\ w}{(u^2 + v^2 + w^2)^{\frac{1}{2}}} \right\} + (\delta - w) \left\{ \frac{H(u^2 + v^2 + 2w^2)}{(u^2 + v^2 + w^2)^{\frac{1}{2}}} \right\} .$$

The system of differential equations in the six variables λ, μ, ψ, κ, φ, and θ adjoint to (3-4) is:

$$\dot{\lambda} = - \kappa \left\{ \frac{a(r^2 - 3x^2)}{r^5} + \frac{abc(r^2 - 5x^2)}{r^7} + \frac{abz(r^2 - 7x^2)}{r^9} \right\} - \varphi \left\{ \frac{a(- 3xy)}{r^5} + \frac{abc(- 5xy)}{r^7} \right.$$

$$\left. + \frac{abz^2(- 7xy)}{r^9} \right\} - \theta \left\{ \frac{a(- 3xy)}{r^5} + \frac{abd(- 5xz)}{r^7} + \frac{abz^2(- 7xz)}{r^9} \right\} ,$$

$$\psi = -\kappa\left\{\frac{a(- 3xz)}{r^5} + \frac{abc(- 5xz)}{r^7} + \frac{abxz(2r^2 - 7z^2)}{r^9}\right\}$$

$$-\varphi\left\{\frac{a(- 3yz)}{r^5} + \frac{abc(-5yz)}{r^7} + \frac{abyz(2r^2 - 7z^2)}{r^9}\right\}$$

$$-\theta\left\{\frac{a(r^2 - 3z^2)}{r^5} + \frac{abd(r^2 - 5z^2)}{r^7} + \frac{abz^2(3r^2 - 7z^2)}{r^9}\right\}$$

$$(3\text{-}5)$$

$$\dot{\mu} = -\kappa\left\{\frac{a(- 3xy)}{r^5} + \frac{abc(- 5xy)}{r^7} + \frac{abz^2(- 7xy)}{r^9}\right\}$$

$$-\varphi\left\{\frac{a(r^2 - 3y^2)}{r^5} + \frac{abc(r^2 - 5y^2)}{r^7} + \frac{abz^2(r^2 - 7y^2)}{r^9}\right\}$$

$$-\theta\left\{\frac{a(- 3yz)}{r^5} + \frac{abd(- 5yz)}{r^7} + \frac{abz^2(- 7yz)}{r^9}\right\}$$

$$\dot{\kappa} = -\lambda - \kappa\left\{\frac{H(2u^2 + v^2 + w^2)}{(u^2 + v^2 + w^2)^{\frac{1}{2}}}\right\} - \varphi\left\{\frac{H\ u\ v}{(u^2 + v^2 + w^2)^{\frac{1}{2}}}\right\} - \theta\left\{\frac{H\ u\ w}{(u^2 + v^2 + w^2)^{\frac{1}{2}}}\right\}$$

$$\dot{\varphi} = -\mu - \kappa\left\{\frac{H\ u\ v}{(u^2 + v^2 + w^2)^{\frac{1}{2}}}\right\} - \varphi\left\{\frac{H(u^2 + 2v^2 + w^2)}{(u^2 + v^2 + w^2)^{\frac{1}{2}}}\right\} - \theta\left\{\frac{H\ v\ w}{(u^2 + v^2 + w^2)^{\frac{1}{2}}}\right\}$$

$$\dot{\theta} = -\psi - \kappa\left\{\frac{H\ u\ w}{(u^2 + v^2 + w^2)^{\frac{1}{2}}}\right\} - \varphi\left\{\frac{H\ v\ w}{(u^2 + v^2 + w^2)^{\frac{1}{2}}}\right\} - \theta\left\{\frac{H(u^2 + v^2 + 2w^2)}{(u^2 + v^2 + w^2)^{\frac{1}{2}}}\right\}$$

In accordance with the prior discussion of errors, a solution to this system of equations is needed. Since first order effects only are being considered, this integration need not be precise. System (3-5) was solved by the Runge-Kutta fourth order method.

B. Truncation Error in Cowell's Method.

Consider the truncation error involved in the computation of an orbit for an earth satellite using Cowell's Method. For numerical integration of equations (3-2) the double integration formula

$$x = (\delta^{-2} + \frac{1}{12} - \frac{1}{240}\delta^{2} + \frac{31}{60480}\delta^{4} - \frac{289}{3628800}\delta^{6}$$

$$+ \frac{317}{28809600}\delta^{8} - \frac{6803477}{2615348736000}\delta^{10} + \ldots) h^{2}\ddot{x}$$

is available. For single integration the formula

$$\dot{x} = (\mu\delta^{-1} - \frac{1}{12}\mu\delta + \frac{11}{720}\mu\delta^{3} - \frac{191}{60480}\mu\delta^{5}$$

$$+ \frac{2497}{3628800}\mu\delta^{7} - \frac{14797}{95800320}\mu\delta^{9} + \ldots) h\ddot{x}$$

is available. It can be shown that two applications of the single integration formula are equivalent to one application of the double integration, since

$$x = (\mu\delta^{-1} - \frac{1}{12}\mu\delta + \frac{11}{720}\mu\delta^{3} + \ldots) h\dot{x}$$

$$= (\mu\delta^{-1} - \frac{1}{12}\mu\delta + \frac{11}{720}\mu\delta^{3} + \ldots)h(\mu\delta^{-1} - \frac{1}{12}\mu\delta + \frac{11}{720}\mu\delta^{3}\ldots)h\ddot{x}$$

$$x = \mu^{2}(\delta^{-1} - \frac{1}{12}\delta + \frac{11}{720}\mu\delta^{3} + \ldots)^{2} h^{2}\ddot{x}$$

$$x = (1 + \frac{\delta^{2}}{4})(\delta^{-2} - \frac{1}{16} + \frac{27}{120}\delta^{2} - \frac{536}{60480}\delta^{4} + \ldots) h^{2}\ddot{x}$$

after simplification, this yields

$$x = (\delta^{-2} + \frac{1}{12} - \frac{1}{240}\delta^{2} + \frac{31}{60480}\delta^{4} + \ldots) h^{2}\ddot{x}$$

After having reduced the system of three second order equations to a system of six first order equations, the integration may be carried out by the formula:

$$x = (\mu\delta^{-1} - \frac{1}{12}\mu\delta + \frac{11}{720}\mu\delta^{3} - \ldots)\, h\dot{x} \tag{3-6}$$

This is an infinite series generally, and a choice of the order of differences to be used must be made in the integration.

The truncation error shall be determined in the following manner. A specific number of terms of (3-6) are chosen (the cases of two and three terms shall be illustrated). Then the formula (3-6) will be expressed in terms of derivatives, so that the numerical integration method can be considered as approximately equivalent to a system of differential equations. This system of differential equations is compared with the starting system of differential equations, so that the residuals $b_i(t)$ due to the truncation error can be determined (Ref. 17).

Now, by definition,

$$\mu = \frac{1}{2}(E^{\frac{1}{2}} + E^{-\frac{1}{2}})$$

$$\delta = E^{\frac{1}{2}} - E^{-\frac{1}{2}}$$

$$\delta^{k} = (E^{\frac{1}{2}} - E^{-\frac{1}{2}})^{k}$$

where

$$E = e^{hD} = 1 + hD + \frac{h^2D^2}{2} + \frac{h^3D^3}{6} + \frac{h^4D^4}{24} + \frac{h^5D^5}{120} + \frac{h^6D^6}{720} + \ldots .$$

Hence

$$\mu = 1 + \frac{h^2D^2}{8} + \frac{h^4D^4}{384} + \frac{h^6D^6}{64(720)} + \ldots ,$$

$$\delta = hD + \frac{h^3D^3}{24} + \frac{h^5D^5}{16(120)} + \frac{h^7D^7}{(128)(7!)} + \ldots ,$$

$$\mu\delta = \frac{1}{2}(E^{\frac{1}{2}} + E^{-\frac{1}{2}})(E^{\frac{1}{2}} - E^{-\frac{1}{2}}) = \frac{1}{2}(D - E^{-1})$$

$$= hD + \frac{h^3D^3}{6} + \frac{h^5D^5}{120} + \frac{h^7D^7}{7!} + \ldots ,$$

$$\mu\delta^{-1} = h^{-1}D^{-1} + \frac{hD}{12} - \frac{h^3D^3}{30(24)} + \frac{h^5D^5}{10(9)(16)(21)} + \ldots ,$$

$$\mu\delta^3 = \frac{1}{2}(E^2 = E^{-2} + 2[E^{-1} - E]) = h^3D^3 + \frac{h^5D^5}{4} + \ldots .$$

Now using only three terms of the difference equation (3-6) gives

$$x \cong (h^{-1}D^{-1} + \frac{hD}{12} - \frac{h^3D^3}{(30)(24)} + \frac{h^5D^5}{(16)(10)(9)(21)} + \ldots$$
$$- \frac{hD}{12} - \frac{h^3D^3}{72} - \frac{h^5D^5}{(12)(120)} + \ldots$$
$$+ \frac{11\ h^3D^3}{720} + \frac{11\ h^5D^5}{(720)(4)} + \ldots)\ h\dot{x}$$

$$\cong (h^{-1}D^{-1} + \frac{191}{60480} h^5D^5 + \ldots)\ h\dot{x}$$

$$\cong (h^{-1}D^{-1} + \frac{191}{60480} h^5D^5 + \ldots)\ hDx ,$$

and

$$Dx \cong (1 + \frac{191}{60480} h^6D^6 + \ldots)\ Dx$$

$$\cong Dx + \frac{191}{60480} h^6D^7x + \ldots ,$$

and from the previous discussion the b_i are approximately

$$b_i(t) = \frac{191}{60480} h^6D^7x .$$

If only two terms in the difference equation (3-6) are used, then

$$Dx = (1 - \frac{11}{720} h^4D^4 + \ldots)\ Dx ,$$

such that

$$b_i(t) = \frac{11}{720} h^4D^5x$$

The truncation error for one time-step is given by:

$$\epsilon_i(t) = \frac{11}{720} h^4 \int_{t-h}^{t} [(D^5 x_i(t)\ \lambda_i(t) + D^5 y_i(t)\ \mu(t) + D^5 z(t)\ \psi(t)$$

$$+ D^5 u(t)\,\kappa(t) + D^5 v(t)\ \varphi(t) + D^5 w(t)\ \theta(t)]dt \ .$$

This integration when extended over all time from t_o to t_n will give the truncation error approximation.

Due to the extensive computation involved in getting $D^7 x_i$, the following analysis will make use of $D^5 x_i$ (i.e., only two terms of (3-6) are used. These fifth derivatives are computed for x, y, z, u, v, and w, and are listed in Appendix C. The third and fourth derivatives are involved here, but will not be given explicitly.

The truncation error analysis proceeds as follows: from the computer program, to be described later, values of the coordinates and velocity components are obtained for times $t = nh$, $n = 0, 1, 2, \ldots$. These are used to evaluate the third, fourth and fifth derivatives at the corresponding times. This requires a small computer program. Furthermore the coefficients of λ, μ, ψ, κ, φ, and θ in the adjoint system are also computed for the same times. Now the adjoint system is solved numerically using Runge-Kutta for the same times and subject to certain initial conditions. In practice there are six sets of initial conditions used. The six initial conditions are:

$$\lambda(t_{100}) = 1; \quad \mu(t_{100}) = \psi(t_{100}) = \kappa(t_{100}) = \varphi(t_{100}) = \theta(t_{100}) = 0 \ ,$$

$$\lambda(t_{100}) = 1; \quad \mu(t_{100}) = \psi(t_{100}) = \kappa(t_{100}) = \varphi(t_{100}) = \theta(t_{100}) = 0 \ ,$$

$$\psi(t_{100}) = 1; \quad \lambda(t_{100}) = \mu(t_{100}) = \kappa(t_{100}) = \varphi(t_{100}) = \theta(t_{100}) = 0 \ ,$$

$$\kappa(t_{100}) = 1; \quad \lambda(t_{100}) = \mu(t_{100}) = \psi(t_{100}) = \varphi(t_{100}) = \theta(t_{100}) = 0 \ ,$$

$$\varphi(t_{100}) = 1; \quad \lambda(t_{100}) = \mu(t_{100}) = \psi(t_{100}) = \kappa(t_{100}) = \theta(t_{100}) = 0 \quad ,$$

$$\theta(t_{100}) = 1; \quad \lambda(t_{100}) = \mu(t_{100}) = \psi(t_{100}) = \kappa(t_{100}) = \varphi(t_{100}) = 0 \quad .$$

Thus six integrations through the adjoint system are accomplished, but it will be observed that these integrations are backward in time from t_{100} to t_o. These integrations could be performed one at a time, but in the actual Runge-Kutta integration program all six were carried along simultaneously. The linearity of the adjoint system simplified the procedure somewhat, and the assumption that the coefficients were constant on the integration interval naturally made the program elementary. The programming was done in a manner which permitted twenty sub-divisions of each time interval.

With these values of the adjoint parameters and the values of the fifth derivatives, another program was written which did a Simpson's Rule integration of the product of corresponding values of these quantities over the range of t_o to t_n for the coordinates only. These three integrals, the errors in the coordinates, as evaluated are:

$$\epsilon_x(t) = \frac{11}{720} h^4 \int_{t_o}^{t_n} [\lambda_1(t)\ D^5x(t) + \mu_1(t)\ d^5y(t) + \psi_1(t)\ D^5z(t)$$

$$+ \kappa_1(t)\ D^5u(t) + \varphi_1(t)_1(t)\ D^5v(t) + \theta_1(t)\ D^5w(t)]\ dt \quad ,$$

$$\epsilon_y(t) = \frac{11}{720} h^4 \int_{t_o}^{t_n} [\lambda_2(t)\ D^5x(t) + \mu_2(t)\ D^5y(t) + \psi_2(t)\ D^5z(t)$$

$$+ \kappa_2(t)\ D^5u(t) + \varphi_2(t)\ D^5v(t) + \theta_2(t)\ D^5w(t)]\ dt$$

$$\epsilon_z(t) = \frac{11}{720} h^4 \int_{t_o}^{t_n} [\lambda_3(t)\ D^5x(t) + \mu_3(t)\ D^5y(t) + \psi_3(t)\ D^5z(t)$$

$$+ \kappa_3(t)\ D^5u(t) + \varphi_3(t)\ D^5v(t) + \theta_3(t)\ D^5w(t)]\ dt \quad ,$$

where the subscripts on the adjoint parameters denote values corresponding to the first, second, and third sets of initial conditions respectively. These truncation errors were computed for $n = 540$ and the results were as follows:

$$\epsilon_x(t) \sim (0.22185059)(10)^{-4}, \quad \epsilon_y(t) \sim (0.34453913)(10)^{-4}, \quad \epsilon_z(t) \sim (0.43451463)(10)^{-4}.$$

In general, the truncation error can be controlled through a choice of the order of differences used in the difference equations and by a choice of the time step h in the integration scheme. If n terms of (3-6) are used, then the order of differences utilized is $2n-3$, and the truncation error is then of the order of h^{2n+1}.

C. Round-Off Errors in Cowell's Method

In order to investigate the accumulation of round-off errors in the integration of differential equations, some assumptions regarding the nature of the integration process must be made. Assume that the six difference tables have been constructed for the initial thirteen values of the functions. The equations to be solved are of the form:

$$\begin{aligned} \dot{x} &= u \\ \dot{y} &= v \\ \dot{z} &= w \\ \dot{u} &= g_1(t, u, v, w, x, y, z) \\ \dot{v} &= g_2(t, u, v, w, x, y, z) \\ \dot{w} &= g_3(t, u, v, w, x, y, z) \end{aligned} \tag{3-7}$$

The numerical integration is accomplished by:

$$x = (\mu\delta^{-1} - \frac{1}{12}\mu\delta + \frac{11}{720}\mu\delta^3 - \ldots)\, hu$$

$$y = (\mu\delta^{-1} - \frac{1}{12}\mu\delta + \frac{11}{720}\mu\delta^3 - \ldots)\, hv$$

$$z = (\mu\delta^{-1} - \frac{1}{12}\mu\delta + \frac{11}{720}\mu\delta^3 - \ldots)\, hw$$

$$u = (\mu\delta^{-1} - \frac{1}{12}\mu\delta + \frac{11}{720}\mu\delta^3 - \ldots) hg_1$$

$$v = (\mu\delta^{-1} - \frac{1}{12}\mu\delta + \frac{11}{720}\mu\delta^3 - \ldots) hg_2$$

$$w = (\mu\delta^{-1} - \frac{1}{12}\mu\delta + \frac{11}{720}\mu\delta^3 - \ldots) hg_3$$

These equations can be written as (Ref. 11):

$$\begin{aligned} x_j &= x_{j-1} + h \sum_{i=1}^{r} \ell_i \\ y_j &= y_{j-1} + h \sum_{i=1}^{r} m_i \\ z_j &= z_{j-1} + h \sum_{i=1}^{r} n_i \\ u_j &= u_{j-1} + h \sum_{i=1}^{r} o_i \\ v_j &= v_{j-1} + h \sum_{i=1}^{r} p_i \\ w_j &= w_{j-1} + h \sum_{i=1}^{r} q_i \end{aligned} \qquad (3\text{-}8)$$

Assuming that h is small enough so that the round-off errors committed in the formation of ℓ_i, m_i, n_i, o_i, p_i, q_i are lost when the multiplication by h is accomplished, there is then only one round-off in computing the next value of x, y, z, u, v, w. To obtain a first approximation to the complete round-off errors, write the equations (3-1) in the form of difference equations and ignore higher powers of h. Then, using the bar for rounded values,

$$\bar{x}_j = \bar{x}_{j-1} + h\bar{u}_{j-1} + \epsilon_j^{(1)} (10)^{-k}$$

$$\bar{y}_j = \bar{y}_{j-1} + h\bar{v}_{j-1} + \epsilon_j^{(2)} (10)^{-k}$$

$$\bar{z}_j = \bar{z}_{j-1} + h\bar{w}_{j-1} + \epsilon_j^{(3)} (10)^{-k}$$

$$\tag{3-9}$$

$$\bar{u}_j = \bar{u}_{j-1} + hg_1(t_{j-1}, \bar{u}_{j-1}, \bar{v}_{j-1}, \bar{w}_{j-1}, \bar{x}_{j-1}, \bar{y}_{j-1}, \bar{z}_{j-1}) + \epsilon_j^{(4)}(10)^{-k}$$

$$\bar{v}_j = \bar{v}_{j-1} + hg_2(t_{j-1}, \bar{u}_{j-1}, \bar{v}_{j-1},\bar{w}_{j-1}, \bar{x}_{j-1}, \bar{y}_{j-1}, \bar{z}_{j-1}) + \epsilon_j^{(5)}(10)^{-k}$$

$$\bar{w}_j = \bar{w}_{j-1} + hg_3(t_{j-1}, \bar{u}_{j-1}, \bar{v}_{j-1},\bar{w}_{j-1}, \bar{x}_{j-1}, \bar{y}_{j-1}, \bar{z}_{j-1}) + \epsilon^{(6)}(10)^{-k}$$

where $\epsilon_j^{(i)}$ denotes the actual round-off at the k^{th} decimal place in the i^{th} variable. For a first approximation, compare these with

$$x_j = x_{j-1} + hu_{j-1}$$

$$y_j = y_{j-1} + hv_{j-1}$$

$$z_j = z_{j-1} + hw_{j-1}$$

$$\tag{3-10}$$

$$u_j = u_{j-1} + hg_1(t_{j-1}, u_{j-1}, v_{j-1}, w_{j-1}, x_{j-1}, y_{j-1}, z_{j-1})$$

$$v_j = v_{j-1} + hg_2(t_{j-1}, u_{j-1}, v_{j-1}, w_{j-1}, x_{j-1}, y_{j-1},z_{j-1})$$

$$w_j = w_{j-1} + hg_3(t_{j-1}, u_{j-1}, v_{j-1}, w_{j-1}, x_{j-1}, y_{j-1}, z_{j-1})$$

to obtain difference equations in the quantities:

$$A_j = x_j - \bar{x}_j$$

$$B_j = y_j - \bar{y}_j$$

$$C_j = z_j - \bar{z}_j$$

$$D_j = u_j - \bar{u}_j$$

$$E_j = v_j - \bar{v}_j$$

$$F_j = w_j - \bar{w}_j$$

By subtracting equations (3-9) from equations (3-10) and expanding functions g_1, g_2, g_3 in a Taylor series through first order terms, then the equations:

$$A_j - A_{j-1} = hD_{j-1} - \epsilon_j^{(1)} (10)^{-k}$$

$$B_j - B_{j-1} = hE_{j-1} - \epsilon_j^{(2)} (10)^{-k}$$

$$C_j - C_{j-1} = hF_{j-1} - \epsilon_j^{(3)} (10)^{-k}$$

(3-11)

$$D_j - D_{j-1} = h\left\{\frac{\partial g_1}{\partial x}A_{j-1} + \frac{\partial g_1}{\partial y}B_{j-1} + \frac{\partial g_1}{\partial z}C_{j-1} + \frac{\partial g_1}{\partial u}D_{j-1} + \frac{\partial g_1}{\partial v}E_{j-1} + \frac{\partial g_1}{\partial w}F_{j-1}\right\} - \epsilon_j^{(4)}(10)^{-k}$$

$$E_j - E_{j-1} = h\left\{\frac{\partial g_2}{\partial x}A_{j-1} + \frac{\partial g_2}{\partial y}B_{j-1} + \frac{\partial g_2}{\partial z}C_{j-1} + \frac{\partial g_2}{\partial u}D_{j-1} + \frac{\partial g_2}{\partial v}E_{j-1} + \frac{\partial g_2}{\partial w}F_{j-1}\right\} - \epsilon_j^{(5)}(10)^{-k}$$

$$F_j - F_{j-1} = h\left\{\frac{\partial g_3}{\partial x}A_{j-1} + \frac{\partial g_3}{\partial y}B_{j-1} + \frac{\partial g_3}{\partial z}C_{j-1} + \frac{\partial g_3}{\partial u}D_{j-1} + \frac{\partial g_3}{\partial v}E_{j-1} + \frac{\partial g_3}{\partial w}F_{j-1}\right\} - \epsilon_j^{(6)}(10)^{-k}$$

are obtained.

The difference equations adjoint to (3-11) are:

$$\lambda_j - \lambda_{j-1} = -h\left\{\frac{\partial g_1}{\partial x}\kappa_j + \frac{\partial g_2}{\partial x}\varphi_j + \frac{\partial g_3}{\partial x}\theta_j\right\}$$

$$\mu_j - \mu_{j-1} = -h\left\{\frac{\partial g_1}{\partial y}\kappa_j + \frac{\partial g_2}{\partial y}\varphi_j + \frac{\partial g_3}{\partial y}\theta_j\right\}$$

$$\psi_j - \psi_{j-1} = -h\left\{\frac{\partial g_1}{\partial z}\kappa_j + \frac{\partial g_2}{\partial z}\varphi_j + \frac{\partial g_3}{\partial y}\theta_j\right\}$$

$$\kappa_j - \kappa_{j-1} = -h\left\{\lambda_j + \frac{\partial g_1}{\partial u}\kappa_j + \frac{\partial g_2}{\partial u}\varphi_j + \frac{\partial g_3}{\partial u}\theta_j\right\} \qquad (3\text{-}12)$$

$$\varphi_j - \varphi_{j-1} = -h\left\{\mu_j + \frac{\partial g_1}{\partial v}\kappa_j + \frac{\partial g_2}{\partial v}\varphi_j + \frac{\partial g_3}{\partial v}\theta_j\right\}$$

$$\theta_j - \theta_{j-1} = -h\left\{\psi_j + \frac{\partial g_1}{\partial w}\kappa_j + \frac{\partial g_2}{\partial w}\varphi_j + \frac{\partial g_3}{\partial w}\theta_j\right\}$$

Multiply the first of equations (3-11) by λ_j, the second by μ_j. the third by ψ_j, the fourth by κ_j, the fifth by φ_j, the sixth by θ_j; multiply the first of equations (3-12) by A_{j-1}, the second by B_{j-1}, the third by C_{j-1}, the fourth by D_{j-1}, the fifth by E_{j-1}, the sixth by F_{j-1}, and then add the results to obtain:

$$\left\{\lambda_j A_j + \mu_j B_j + \psi_j C_j + \kappa_j D_j + \varphi_j E_j + \theta_j F_j\right\} \qquad (3\text{-}13)$$

$$- \left\{\lambda_{j-1} A_{j-1} + \mu_{j-1} B_{j-1} + \psi_{j-1} C_{j-1} + \kappa_{j-1} + \varphi_{j-1} E_{j-1} + \theta_{j-1} F_{j-1}\right\}$$

$$= -(10)^{-k}\left\{\epsilon_j^{(1)}\lambda_j + \epsilon_j^{(2)}\mu_j + \epsilon_j^{(3)}\psi_j + \epsilon_j^{(4)}\kappa_{j-1} + \epsilon_j^{(5)}\varphi_j + \epsilon_j^{(6)}\theta_j\right\}$$

Summing up for $j = 1, 2, \ldots, n$ yields:

$$\left\{\lambda(t_n)\ A_n + \mu(t_n)\ B_n + \psi(t_n)\ C_n + \kappa(t_n)\ D_n + \varphi(t_n)\ E_n + \theta(t_n)\ F_n\right\}$$

$$- \left\{\lambda(t_o)\ A_o + \mu(t_o)\ B_o + \psi(t_o)\ C_o + \kappa(t_o)\ D_o + \varphi(t_o)\ E_o + \theta(t_o)\ F_o\right\} \quad \text{(3-1}$$

$$= -(10)^{-k} \sum_{j=1}^{n} \left\{\epsilon_j^{(1)}\lambda(t_j) + \epsilon_j^{(2)}\mu(t_j) + \epsilon_j^{(3)}\psi(t_j) + \epsilon_j^{(4)}\kappa(t_j) + \epsilon_j^{(5)}\varphi(t_j) + \epsilon_j^{(6)}\theta(t_j)\right.$$

If the solution of the adjoint system (3-6) is chosen so that $\lambda(t_n) = 1$, $\mu(t_n) = \psi(t_n) = \kappa(t_n) = \varphi(t_n) = \theta(t_n) = 0$, then

$$A_n = -(10)^{-k} \sum_{j=o}^{n} \left\{\epsilon_j^{(1)}\lambda(t_j) + \epsilon_j^{(2)}\mu(t_j) + \epsilon_j^{(3)}\psi(t_j) + \epsilon_j^{(4)}\kappa(t_j) + \epsilon_j^{(5)}\varphi(t_j) + \epsilon_j^{(6)}\theta(t_j)\right.$$

If the solution of the adjoint system (3-6) is chosen so that $\lambda(t_n) = 0$, $\mu(t_n) = 1$; $\psi(t_n) = \kappa(t_n) = \varphi(t_n) = \theta(t_n) = 0$, then

$$B_n = -(10)^{-k} \sum_{j=o}^{n} \left\{\epsilon_j^{(1)}\lambda(t_j) + \epsilon_j^{(2)}\mu(t_j) + \epsilon_j^{(3)}\psi(t_j) + \epsilon_j^{(4)}\kappa(t_j) + \epsilon_j^{(5)}\varphi(t_j) + \epsilon_j^{(6)}\theta(t_j)\right\}$$

In similar fashion, equations for C_n, D_n, E_n and F_n may be obtainec

Since $\epsilon_j^{(i)}$ $(i = 1,2,3,4,5,6)$ are assumed to be uniformly distributed in the range $-\frac{1}{2} \leq \epsilon_j^{(1)} \leq \frac{1}{2}$, then $E(\epsilon_j^{(i)}) = 0$, and $\sigma^2(\epsilon_j^{(i)}) = \frac{1}{12}$, where $E(\epsilon_j^{(i)}$ denotes the expected value of $\epsilon_j^{(i)}$, and $\sigma^2(\epsilon_j^{(i)})$ the variance.

Accordingly,

$$\sigma^2(A_n) = \frac{(10)^{-2k}}{12h} \sum_{j=o}^{n} \left\{\lambda_j^2(t_n) + \mu_j^2(t_j) + \psi_j^2(t_j) + \kappa_j^2(t_j) + \varphi_j^2(t_j) + \theta_j^2(t_j)\right\} \text{c}$$

$$\sigma^2(A_n) = \frac{(10)^{-2k}}{12h} \int_{t_o}^{t_n} [\lambda^2(t) + \mu^2(t) + \psi^2(t) + \kappa^2(t) + \varphi^2(t) + \theta^2(t)]\, dt ,$$

with similar expressions for $\sigma^2(B_n)$, $\sigma^2(C_n)$, $\sigma^2(D_n)$, $\sigma^2(E_n)$, and $\sigma^2(F_n)$.

Now $P\{|A_n| \geqq \epsilon\} = P\{A_n^2 \geqq \epsilon^2\}$, where $P\{|A_n| \geqq \epsilon\}$ stands for the probability that $|A_n|$ be greater than or equal to ϵ . By means of the Bienaymé-Tchebycheff inequality, $P\{|A_n| \geqq \epsilon\} = P\{A_n^2 \geqq \epsilon^2\} \leqq \frac{E(A_n^2)}{\epsilon^2} = \frac{\sigma^2(A_n)}{\epsilon^2}$, since $E(A_n) = 0$ (Ref. 15, p. 182).

Hence,

$$P\{|A_n| \geqq \epsilon\} \leqq \frac{(10)^{-2k} \int_{t_o}^{t_n} [\lambda^2(t) + \mu^2(t) + \psi^2(t) + \kappa^2(t) + \varphi^2(t) + \theta^2(t)]\, dt}{12h\epsilon^2} ,$$

with similar expressions for B_n, C_n, D_n, E_n, and F_n . If ϵ is chosen so that $P\{|A_n| \geqq \epsilon\} \leqq 0.01$, than $|A_n| < \epsilon$ 99 per cent of the time. The results for the case $n = 540$ are:

$$|A_n| \leqq (6.636)\ (10)^{-9}$$

$$|B_n| \leqq (6.414)\ (10)^{-9}$$

$$|C_n| \leqq (8.429)\ (10)^{-9}$$

These results are valid 99 per cent of the time.

D. Propagated Error:

The propagated errors in terms of the estimates of the errors in the initial values of the coordinates and velocities are given by:

$$E_x = e_1\lambda_1 + e_2\mu_1 + e_3\psi_1 + e_4\kappa_1 + e_5\varphi_1 + e_6\theta_1$$

$$E_y = e_1\lambda_2 + e_2\mu_2 + e_3\psi_2 + e_4\kappa_2 + e_5\varphi_2 + e_6\theta_2$$

$$E_z = e_1\lambda_3 + e_2\mu_3 + e_3\psi_3 + e_4\kappa_3 + e_5\varphi_3 + e_6\theta_3$$

$$E_u = e_1\lambda_4 + e_2\mu_4 + e_3\psi_4 + e_4\kappa_4 + e_5\varphi_4 + e_6\theta_4$$

$$E_v = e_1\lambda_5 + e_2\mu_5 + e_3\psi_5 + e_4\kappa_5 + e_5\varphi_5 + e_6\theta_5$$

$$E_w = e_1\lambda_6 + e_2\mu_6 + e_3\psi_6 + e_4\kappa_6 + e_5\varphi_6 + e_6\theta_6$$

where e_1, e_2, e_3, e_4, e_5, e_6 are the estimates of the errors in the initial values of x, y, z, u, v, w. The six sets of values of the adjoint parameters correspond to six different solutions of the adjoint equations, i.e., the set λ_1, μ_1, . . ., θ_1 corresponds to the solution with conditions $\lambda(t_n) = 1$, $\mu(t_n) = . . . = \theta(t_n) = 0$ evaluated at $t = t_o$, and similarly for the other sets.

4. The Test Program for Orbit of Satellite.

Relative to the equations of motion (3-1) the following input values were known: x_o, y_o, z_o, $\dot{x}_o$, $\dot{y}_o$, $\dot{z}_o$ at t_o and k, $M_\oplus$, J, R_e, θ, C_d. Compute the following quantities:

$$r_o^2 = x_o^2 + y_o^2 + z_o^2 \qquad (4\text{-}1)$$

$$v_o^2 = \dot{x}_o^2 + \dot{y}_o^2 + \dot{z}_o^2 \qquad (4\text{-}2)$$

$$\left(\frac{\partial U}{\partial x}\right)_o = -\frac{k^2 M_\oplus}{r_o^2}\left\{1 - \frac{5JR_e^2}{r_o^2}\left[\left(\frac{z_o}{r_o}\right)^2 - \frac{1}{5}\right]\right\}\frac{x_o}{r_o} \quad ,$$

$$\left(\frac{\partial U}{\partial y}\right)_o = -\frac{k^2 M_\oplus}{r_o^2}\left\{1 - \frac{5JR_e^2}{r_o^2}\left[\left(\frac{z_o}{r_o}\right)^2 - \frac{1}{5}\right]\right\}\frac{y_o}{r_o} \quad ,$$

$$\left(\frac{\partial U}{\partial z}\right)_o = -\frac{k^2 M_\oplus}{r_o^2}\left\{1 - \frac{5JR_e^2}{r_o^2}\left[\left(\frac{z_o}{r_o}\right)^2 - \frac{3}{5}\right]\right\}\frac{z_o}{r_o} \quad ,$$

$$D_{x_o} = C_d \sigma_o \theta \, V_o \dot{x}_o \quad ,$$

$$D_{y_o} = C_d \sigma_o \theta \, V_o \dot{y}_o \quad ,$$

$$D_{z_o} = C_d \sigma_o \theta \, V_o \dot{z}_o \quad ,$$

where $U(x, y, z)$ is the potential function for the two-body problem, D_x, D_y, and D_z are the acceleration components due to drag, and the zero subscript indicates that the quantities are to be computed at time t_o. The quantity σ_o is to be determined from the following formulae (Ref. 16 pp. 20).

$$T = 5.80441045 \times 10^4 \left[\frac{h}{1 + \frac{h}{0.98747414}} \right] - 603.318 , \qquad (4\text{-}3)$$

$$\log_{10} \sigma_o = 5.588 - 4.79 \log_{10} T , \qquad (4\text{-}4)$$

where $h = r_o - R_m$ and R_m is the mean radius of the earth. Next compute the following quantities:

$$X_o = \left(\frac{\partial U}{\partial x}\right)_o + D_{x_o} , \qquad (4\text{-}5)$$

$$Y_o = \left(\frac{\partial U}{\partial y}\right)_o + D_{y_o} , \qquad (4\text{-}6)$$

$$Z_o = \left(\frac{\partial U}{\partial z}\right)_o + D_{z_o} . \qquad (4\text{-}7)$$

The initial difference table is constructed at this point. These computations extend through equation (4-7) . Let the interval of integration be $\Delta t \sim 1^m = 0.0007$ mean solar day. Compute the following approximate second and first sums in the integration scheme (central differences):

$$''X_o = \frac{x_o}{(\Delta t)^2} - \frac{1}{12} X_o , \qquad (4\text{-}8)$$

$$'X_{-\frac{1}{2}} = \frac{\dot{x}_o}{\Delta t} - \frac{1}{2} X_o , \qquad (4\text{-}9)$$

$$'X_{\frac{1}{2}} = \frac{\dot{x}_o}{\Delta t} + \frac{1}{2} X_o' . \qquad (4\text{-}10)$$

The calculations for Y and Z are made by formulae similar to (4-8), (4-9) and (4-10) . All three calculations must be made concurrently.

Next, compute the following:

$$
\begin{aligned}
x_{-6} &= (\Delta t)^2 \,[\,''X_o - 6\,'X_{-\frac{1}{2}} + \tfrac{181}{12}\,X_o\,] \,, \\
x_{-5} &= (\Delta t)^2 \,[\,''X_o - 5\,'X_{-\frac{1}{2}} + \tfrac{121}{12}\,X_o\,] \,, \\
x_{-4} &= (\Delta t)^2 \,[\,''X_o - 4\,'X_{-\frac{1}{2}} + \tfrac{73}{12}\,X_o\,] \,, \\
x_{-3} &= (\Delta t)^2 \,[\,''X_o - 3\,'X_{-\frac{1}{2}} + \tfrac{37}{12}\,X_o\,] \,, \\
x_{-2} &= (\Delta t)^2 \,[\,''X_o - 2\,'X_{-\frac{1}{2}} + \tfrac{13}{12}\,X_o\,] \,, \\
x_{-1} &= (\Delta t)^2 \,[\,''X_o - \,'X_{-\frac{1}{2}} + \tfrac{1}{12}\,X_o\,] \,, \\
x_{1} &= (\Delta t)^2 \,[\,''X_o + \,'X_{\frac{1}{2}} + \tfrac{1}{12}\,X_o\,] \,, \\
x_{2} &= (\Delta t)^2 \,[\,''X_o + 2\,'X_{\frac{1}{2}} + \tfrac{13}{12}\,X_o\,] \,, \\
x_{3} &= (\Delta t)^2 \,[\,''X_o + 3\,'X_{\frac{1}{2}} + \tfrac{37}{12}\,X_o\,] \,, \\
x_{4} &= (\Delta t)^2 \,[\,''X_o + 4\,'X_{\frac{1}{2}} + \tfrac{73}{12}\,X_o\,] \,, \\
x_{5} &= (\Delta t)^2 \,[\,''X_o + 5\,'X_{\frac{1}{2}} + \tfrac{121}{12}\,X_o\,] \,, \\
x_{6} &= (\Delta t)^2 \,[\,''X_o + 6\,'X_{\frac{1}{2}} + \tfrac{181}{12}\,X_o\,] \,,
\end{aligned}
\qquad (4\text{-}11)
$$

$$
\begin{aligned}
\dot{x}_{-6} &= \Delta t\,[\,{}'X_{-\frac{1}{2}} - \tfrac{11}{2} X_o] \,, \\
\dot{x}_{-5} &= \Delta t\,[\,{}'X_{-\frac{1}{2}} - \tfrac{9}{2} X_o] \,, \\
\dot{x}_{-4} &= \Delta t\,[\,{}'X_{-\frac{1}{2}} - \tfrac{7}{2} X_o] \,, \\
\dot{x}_{-3} &= \Delta t[\,{}'X_{-\frac{1}{2}} - \tfrac{5}{2} X_o] \,, \\
\dot{x}_{-2} &= \Delta t[\,{}'X_{-\frac{1}{2}} - \tfrac{3}{2} X_o] \,, \\
\dot{x}_{-1} &= \Delta t[\,{}'X_{-\frac{1}{2}} - \tfrac{1}{2} X_o] \,, \\
\dot{x}_{1} &= \Delta t[\,{}'X_{\frac{1}{2}} + \tfrac{1}{2} X_o] \,, \\
\dot{x}_{2} &= \Delta t[\,{}'X_{\frac{1}{2}} + \tfrac{3}{2} X_o] \,, \\
\dot{x}_{3} &= \Delta t[\,{}'X_{\frac{1}{2}} + \tfrac{5}{2} X_o] \,, \\
\dot{x}_{4} &= \Delta t[\,{}'X_{\frac{1}{2}} + \tfrac{7}{2} X_o] \,, \\
\dot{x}_{5} &= \Delta t[\,{}'X_{\frac{1}{2}} + \tfrac{9}{2} X_o] \,, \\
\dot{x}_{6} &= \Delta t[\,{}'X_{\frac{1}{2}} + \tfrac{11}{2} X_o] \,.
\end{aligned}
\qquad (4\text{-}12)
$$

The values for $\dot{x}_i$ and x_i computed above and similarly computed values for $\dot{y}_i$, y_i, $\dot{z}_i$ and z_i are substituted into equations (4-1) through (4-7) to obtain values for X_i, Y_i, Z_i, $-6 \leqq i \leqq 6$ and $i \neq 0$. These values of X_i, Y_i, Z_i are now used to form three difference tables; the one for X is given as follows:

		x_{-6}										
			$\delta x_{-11/2}$									
		x_{-5}		$\delta^2 x_{-5}$								
			$\delta x_{-9/2}$		$\delta^3 x_{-9/2}$							
		x_{-4}		$\delta^2 x_{-4}$		$\delta^4 x_{-4}$						
			$\delta x_{-7/2}$		$\delta^3 x_{-7/2}$		$\delta^5 x_{-7/2}$					
		x_{-3}		$\delta^2 x_{-3}$		$\delta^4 x_{-3}$		$\delta^6 x_{-3}$				
			$\delta x_{-5/2}$		$\delta^3 x_{-5/2}$		$\delta^5 x_{-5/2}$		$\delta^7 x_{-5/2}$			
		x_{-2}		$\delta^2 x_{-2}$		$\delta^4 x_{-2}$		$\delta^6 x_{-2}$		$\delta^8 x_{-2}$		
			$\delta x_{-3/2}$		$\delta^3 x_{-3/2}$		$\delta^5 x_{-3/2}$		$\delta^7 x_{-3/2}$		$\delta^9 x_{-3/2}$	
		x_{-1}		$\delta^2 x_{-1}$		$\delta^4 x_{-1}$		$\delta^6 x_{-1}$		$\delta^8 x_{-1}$		$\delta^{10} x_{-1}$
	$'x_{-1/2}$	$x_{-1/2}$	$\delta x_{-1/2}$		$\delta^3 x_{-1/2}$		$\delta^5 x_{-1/2}$		$\delta^7 x_{-1/2}$		$\delta^9 x_{-1/2}$	
$''x_0$		x_0		$\delta^2 x_0$		$\delta^4 x_0$		$\delta^6 x_0$		$\delta^8 x_0$		$\delta^{10} x_0$
	$'x_{1/2}$	$x_{1/2}$	$\delta x_{1/2}$		$\delta^3 x_{1/2}$		$\delta^5 x_{1/2}$		$\delta^7 x_{1/2}$		$\delta^9 x_{1/2}$	
		x_1		$\delta^2 x_1$		$\delta^4 x_1$		$\delta^6 x_1$		$\delta^8 x_1$		$\delta^{10} x_1$
			$\delta x_{3/2}$		$\delta^3 x_{3/2}$		$\delta^5 x_{3/2}$		$\delta^7 x_{3/2}$		$\delta^9 x_{3/2}$	
		x_2		$\delta^2 x_2$		$\delta^4 x_2$		$\delta^6 x_2$		$\delta^8 x_2$		
			$\delta x_{5/2}$		$\delta^3 x_{5/2}$		$\delta^5 x_{5/2}$		$\delta^7 x_{5/2}$			
		x_3		$\delta^2 x_3$		$\delta^4 x_3$		$\delta^6 x_3$				
			$\delta x_{7/2}$		$\delta^3 x_{7/2}$		$\delta^5 x_{7/2}$					
		x_4		$\delta^2 x_4$		$\delta^4 x_4$						
			$\delta x_{9/2}$		$\delta^3 x_{9/2}$							
		x_5		$\delta^2 x_5$								
			$\delta x_{11/2}$									
		x_6										

(4-13)

Using the differences from the difference tables, calculate $"X_o$, $'X_o$, $'X_{-\frac{1}{2}}$, $'X_{\frac{1}{2}}$ by the following formulae:

$$"X_o = \frac{x_o}{(\Delta t)^2} - \frac{1}{12} X_o + \frac{1}{240} \delta^2 X_o - \frac{31}{60480} \delta^4 X_o$$
$$+ \frac{289}{3628800} \delta^6 X_o - \frac{317}{22809600} \delta^8 X_o + \frac{6803477}{261534873600} \delta^{10} X_o \,, \tag{4-14}$$

$$'X_o = \frac{\dot{x}_o}{\Delta t} + \frac{1}{12} \mu\delta X_o - \frac{11}{720} \mu\delta^3 X_o + \frac{191}{60480} \mu\delta^5 X_o$$
$$- \frac{2497}{3628800} \mu\delta^7 X_o + \frac{14797}{95800320} \mu\delta^9 X_o \,, \tag{4-15}$$

$$'X_{-\frac{1}{2}} = {}'X_o - \frac{1}{2} X_o \,, \tag{4-16}$$

$$'X_{\frac{1}{2}} = {}'X_o + \frac{1}{2} X_o \,. \tag{4-17}$$

These calculations yield improved values of $"X_o$, $'X_{-\frac{1}{2}}$ and $'X_{\frac{1}{2}}$. Use these improved values to recalculate all X_i, $-6 \leqq i \leqq 6$ for $i \neq 0$, and then re-form the difference tables. Again recalculate $"X_o$, $'X_o$, $'X_{-\frac{1}{2}}$, and $'X_{\frac{1}{2}}$ by (4-14) through (4-17). This iterative process is continued until the four values $"X_o$, $'X_o$, $'X_{-\frac{1}{2}}$, and $'X_{\frac{1}{2}}$ have settled down, i.e., until two successive determinations of these values differ by less than some prescribed amount. When convergence has taken place, the difference tables are ready for use.

Before proceeding with the step-by-step integration scheme, the difference table (4-13) must be extended to be:

$''x_{-7}$												
	$'x_{-13/2}$											
$''x_{-6}$		x_{-6}										
	$'x_{-11/2}$		$\delta x_{-11/2}$									
$''x_{-5}$		x_{-5}		$\delta^2 x_{-5}$								
	$'x_{-9/2}$		$\delta x_{-9/2}$		$\delta^3 x_{-9/2}$							
$''x_{-4}$		x_{-4}		$\delta^2 x_{-4}$		$\delta^4 x_{-4}$						
	$'x_{-7/2}$		$\delta x_{-7/2}$		$\delta^3 x_{-7/2}$		$\delta^5 x_{-7/2}$					
$''x_{-3}$		x_{-3}		$\delta^2 x_{-3}$		$\delta^4 x_{-3}$		$\delta^6 x_{-3}$				
	$'x_{-5/2}$		$\delta x_{-5/2}$		$\delta^3 x_{-5/2}$		$\delta^5 x_{-5/2}$		$\delta^7 x_{-5/2}$			
$''x_{-2}$		x_{-2}		$\delta^2 x_{-2}$		$\delta^4 x_{-2}$		$\delta^6 x_{-2}$		$\delta^8 x_{-2}$		
	$'x_{-3/2}$		$\delta x_{-3/2}$		$\delta^3 x_{-3/2}$		$\delta^5 x_{-3/2}$		$\delta^7 x_{-3/2}$		$\delta^9 x_{-3/2}$	
$''x_{-1}$		x_{-1}		$\delta^2 x_{-1}$		$\delta^4 x_{-1}$		$\delta^6 x_{-1}$		$\delta^8 x_{-1}$		$\delta^{10} x_{-1}$
	$'x_{-1/2}$		$\delta x_{-1/2}$		$\delta^3 x_{-1/2}$		$\delta^5 x_{-1/2}$		$\delta^7 x_{-1/2}$		$\delta^9 x_{-1/2}$	
$''x_0$		x_0		$\delta^2 x_0$		$\delta^4 x_0$		$\delta^6 x_0$		$\delta^8 x_0$		$\delta^{10} x_0$
	$'x_{1/2}$		$\delta x_{1/2}$		$\delta^3 x_{1/2}$		$\delta^5 x_{1/2}$		$\delta^7 x_{1/2}$		$\delta^9 x_{1/2}$	
$''x_1$		x_1		$\delta^2 x_1$		$\delta^4 x_1$		$\delta^6 x_1$		$\delta^8 x_1$		$\delta^{10} x_1$
	$'x_{3/2}$		$\delta x_{3/2}$		$\delta^3 x_{3/2}$		$\delta^5 x_{3/2}$		$\delta^7 x_{3/2}$		$\delta^9 x_{3/2}$	
$''x_2$		x_2		$\delta^2 x_2$		$\delta^4 x_2$		$\delta^6 x_2$		$\delta^8 x_2$		$\delta^{10} x_2$
	$'x_{5/2}$		$\delta x_{5/2}$		$\delta^3 x_{5/2}$		$\delta^5 x_{5/2}$		$\delta^7 x_{5/2}$		$\delta^9 x_{5/2}$	
$''x_3$		x_3		$\delta^2 x_3$		$\delta^4 x_3$		$\delta^6 x_3$		$\delta^8 x_3$		$\delta^{10} x_3$
	$'x_{7/2}$		$\delta x_{7/2}$		$\delta^3 x_{7/2}$		$\delta^5 x_{7/2}$		$\delta^7 x_{7/2}$		$\delta^9 x_{7/2}$	
$''x_4$		x_4		$\delta^2 x_4$		$\delta^4 x_4$		$\delta^6 x_4$		$\delta^8 x_4$		$\delta^{10} x_4$
	$'x_{9/2}$		$\delta x_{9/2}$		$\delta^3 x_{9/2}$		$\delta^5 x_{9/2}$		$\delta^7 x_{9/2}$		$\delta^9 x_{9/2}$	
$''x_5$		x_5		$\delta^2 x_5$		$\delta^4 x_5$		$\delta^6 x_5$		$\delta^8 x_5$		$\delta^{10} x_5$
	$'x_{11/2}$		$\delta x_{11/2}$		$\delta^3 x_{11/2}$		$\delta^5 x_{11/2}$		$\delta^7 x_{11/2}$		$\delta^9 x_{11/2}$	
$''x_6$		x_6		$\delta^2 x_6$		$\delta^4 x_6$		$\delta^6 x_6$		$\delta^8 x_6$		$\delta^{10} x_6$
	$'x_{13/2}$		$\delta x_{13/2}$		$\delta^3 x_{13/2}$		$\delta^5 x_{13/2}$		$\delta^7 x_{13/2}$		$\delta^9 x_{13/2}$	
$''x_7$		x_7	$\mu\delta x_7$	$\delta^2 x_7$	$\mu\delta^3 x_7$	$\delta^4 x_7$	$\mu\delta^5 x_7$	$\delta^6 x_7$	$\mu\delta^7 x_7$	$\delta^8 x_7$	$\mu\delta^9 x_7$	$\delta^{10} x_7$
			$\delta x_{15/2}$		$\delta^3 x_{15/2}$		$\delta^5 x_{15/2}$		$\delta^7 x_{15/2}$		$\delta^9 x_{15/2}$	

(4-18)

The extension on the left is accomplished as follows:

From definitions of central differences,

$$'X_{-\frac{3}{2}} = 'X_{-\frac{1}{2}} - X_{-1}, \quad 'X_{-\frac{5}{2}} = 'X_{-\frac{3}{2}} - X_{-2}, \quad \ldots, \quad 'X_{-\frac{13}{2}} = 'X_{-\frac{11}{2}} - X_{-6},$$

$$'X_{\frac{3}{2}} = X_1 + 'X_{\frac{1}{2}}, \quad 'X_{\frac{5}{2}} = X_2 + 'X_{\frac{3}{2}}, \quad \ldots, \quad 'X_{\frac{13}{2}} = X_6 + 'X_{\frac{11}{2}},$$

$$''X_{-1} = ''X_o - 'X_{-\frac{1}{2}}, \quad ''X_{-2} = ''X_{-1} - 'X_{-\frac{3}{2}}, \quad \ldots, \quad ''X_{-7} = ''X_{-6} - ''X_{-\frac{13}{2}},$$

$$''X_1 = 'X_{\frac{1}{2}} + ''X_o, \quad ''X_2 = 'X_{\frac{3}{2}} + ''X_1, \quad \ldots, \quad ''X_7 = 'X_{\frac{13}{2}} + ''X_6 .$$

The extension on the right is accomplished as follows:

First, in order to make the tenth differences constant, set

$$\delta^{10}X_2 = \delta^{10}X_{\frac{3}{2}} = \ldots = \delta^{10}X_7 = \delta^{10}X_1 .$$

Then, from the definitions of central differences:

$$\delta^9 X_{\frac{5}{2}} = \delta^9 X_{\frac{3}{2}} + \delta^{10}X_2, \quad \delta^9 X_{\frac{7}{2}} = \delta^9 X_{\frac{5}{2}} + \delta^{10}X_3, \quad \ldots, \quad \delta^9 X_{\frac{15}{2}} = \delta^9 X_{\frac{13}{2}} + \delta^{10}X_7,$$

$$\delta^8 X_3 = \delta^8 X_2 + \delta^9 X_{\frac{5}{2}}, \quad \delta^8 X_4 = \delta^8 X_3 + \delta^9 X_{\frac{7}{2}}, \quad \ldots, \quad \delta^8 X_7 = \delta^8 X_6 + \delta^9 X_{\frac{13}{2}}, \quad \ldots,$$

$$X_7 = X_6 + \delta X_{\frac{13}{2}}$$

Predicted values for x_7 and $\dot{x}_7$ are obtained from (4-14) and (4-15) and are:

$$x_o = (\Delta t)^2 \left[{}''X_o + \frac{1}{12} X_o - \frac{1}{240} \delta^2 X_o + \frac{31}{60480} \delta^4 X_o \right.$$

$$\left. - \frac{289}{3628800} \delta^6 X_o + \frac{317}{22809600} \delta^8 X_o - \frac{6803477}{261534873600} \delta^{10} X_o \right] \tag{4-19}$$

$$\dot{x}_o = (\Delta t) \left[{}'X_o - \frac{1}{12} \mu\delta X_o + \frac{11}{720} \mu\delta^3 X_o - \frac{191}{60480} \mu\delta^5 X_o \right.$$

$$\left. + \frac{2497}{3628800} \mu\delta^7 X_o - \frac{14797}{95800320} \mu\delta^9 X \right].$$

However, instead of utilizing the difference tables to calculate predicted values of x_i and $\dot{x}_i$ these values are computed by means of the following formulae (see Appendix D):

x_i(predictor)

$$\begin{aligned} = (\Delta t)^2 [{}''X_i &+ 0.766936626\ X_{i-1} \\ &- 3.525812185\ X_{i-2} \\ &+ 10.1575288\ X_{i-3} \\ &- 19.87096838\ X_{i-4} \\ &+ 27.44087718\ X_{i-5} \\ &- 27.18762117\ X_{i-6} \\ &+ 19.29048178\ X_{i-7} \\ &- 9.59681704\ X_{i-8} \\ &+ 3.186400889\ X_{i-9} \\ &- 0.6352768226 X_{i-10} \\ &+ 0.05760362585 X_{i-11}] \end{aligned} \tag{4-20}$$

$\dot{x}_i$(predictor)

$$\begin{aligned} = (\Delta T) [{}'X_{i-\frac{1}{2}} &+ 3.726253942\ X_{i-1} \\ &- 16.55959267\ X_{i-2} \\ &+ 47.77560266\ X_{i-3} \\ &- 93.74262788\ X_{i-4} \\ &+ 129.77950265\ X_{i-5} \\ &- 129.0120816\ X_{i-6} \\ &+ 91.5353561\ X_{i-7} \\ &- 45.58936194\ X_{i-8} \\ &+ 15.15063112\ X_{i-9} \\ &- 3.022844997\ X_{i-10} \\ &+ 0.274265540\ X_{i-11}] \end{aligned} \tag{4-21}$$

These formulae are derived directly from the difference table (4-18) and formulae (4-19) and (4-16). By using these formulae the difference tables need not be stored. Only the quantities X_{i-1}, X_{i-2}, . . ., X_{i-11}, $''X_i$ and $'X_{i-\frac{1}{2}}$ are needed. With the x_i(predicted) and $\dot{x}_i$(predicted), calculate X_i from the equations of motion. This new X_i is now used in the following two equations to improve the values of x_i and $\dot{x}_i$:

x_i(corrector)

$$
\begin{aligned}
= (\Delta t)^2 [''X_i &+ 0.05760362585\, X_i \\
&+ 0.1332967418\, X_{i-1} \\
&- 0.357612765\, X_{i-2} \\
&+ 0.652930536\, X_{i-3} \\
&- 0.861771849\, X_{i-4} \\
&+ 0.828002030\, X_{i-5} \\
&- 0.574745365\, X_{i-6} \\
&+ 0.2812852632\, X_{i-7} \\
&- 0.0922187772\, X_{i-8} \\
&+ 0.01820146864\, X_{i-9} \\
&- 0.001636938289 X_{i-10}\,]
\end{aligned}
\tag{4-22}
$$

$\dot{x}_i$(corrector)

$$
\begin{aligned}
= \Delta t\, ['X_{i-\frac{1}{2}} &+ 0.779208129\, X_i \\
&+ 0.709332999\, X_{i-1} \\
&- 1.474887958\, X_{i-2} \\
&+ 2.52178854\, X_{i-3} \\
&- 3.23963332\, X_{i-4} \\
&+ 3.06882316\, X_{i-5} \\
&- 2.11191793\, X_{i-6} \\
&+ 1.02767434\, X_{i-7} \\
&- 0.335547736\, X_{i-8} \\
&+ 0.066064136\, X_{i-9} \\
&- 0.00592405635 X_{i-10}\,].
\end{aligned}
\tag{4-23}
$$

Using the corrected x_i and $\dot{x}_i$, X_i is recomputed from the equations of motion. If the difference between this X_i and the former X_i is less than ϵ (a preassigned quantity), then the integration moves forward. If the difference is not less than ϵ, then iterations are performed using the corrector formulae only until the X_i has settled down.

Before continuing the integration process to find x_{i+1} and $\dot{x}_{i+1}$, it is necessary to compute $'X_{i+\frac{1}{2}}$ and $''X_{i+1}$ by means of

$$'X_{i+\frac{1}{2}} = X_i + 'X_{i-\frac{1}{2}}, \quad \text{and} \quad ''X_{i+1} = ''X_i + 'X_{i+\frac{1}{2}} .$$

After this is done, the predictor formulae are used to obtain x_{i+1}(predicted) and $\dot{x}_{i+1}$(predicted), and then the corrector formulae are applied as before to finally determine x_{i+1} and $\dot{x}_{i+1}$.

Testing for halving or doubling the interval is accomplished by recomputing x_{i-5} using the equation:

$$\begin{aligned} x_{i-5} = (\Delta t)^2[''X_{i-5} & - 0.000002601365128\ X_i \\ & + 0.00003991130840\ X_{i-1} \\ & - 0.0003078833402\ X_{i-2} \\ & + 0.001691708262\ X_{i-3} \\ & - 0.008736096486\ X_{i-4} \\ & + 0.09796325654\ X_{i-5} \\ & - 0.00873609648\ X_{i-6} \\ & + 0.001691708267\ X_{i-7} \\ & - 0.0003078832502\ X_{i-8} \\ & + 0.00003991130840\ X_{i-9} \\ & - 0.000002601365128\ X_{i-10}] . \end{aligned}$$

If the x_{i-5} (new) and x_{i-5} (old) do not compare to the number of significant digits decided upon, halve the time interval. If x_{i-5} (new) and x_{i-5} (old) compare for N_3 (arbitrarily set) consecutive time steps, then double the interval.

The following options were made:

(a) Output. Print out occurs at the end of N_1 (arbitrarily set) iterations in the table-forming stage. A special printout of t and Δt occurs when the interval is changed.

(b) Equations (4-1) through (4-7) were programmed as a fixed location subroutine.

(c) The various convergence tests were made on relative comparison of the two values rather than a comparison with a preassigned ϵ .

(d) The test for doubling or halving the interval can be eliminated.

5 Conclusions and Recommendations.

The process of definitive satellite orbit determination involves three rather general phases:

(1) The determination of initial elements or initial positions on the orbit.

(2) A procedure for calculating future positions on the orbit.

(3) A correction scheme for improving orbit positions determined in (1) and (2) above.

Specifically, phase (1) deals with the preliminary description of the orbit corresponding to some epoch. These initial values of the orbital elements or initial positions on the orbit may be obtained from visual, photographic, or radar observations. The errors in the initial conditions profoundly influence the computed positions, and are caused by errors in the measurements and errors arising from the methods of computation employed to convert observed data into desired initial conditions. Careful consideration should be given to the selection of the quantities to be measured and to the method of computation in order to minimize the errors in the initial conditions. Since data in the vicinity of launch from any source would be helpful, any group engaged in orbit computation should be prepared to utilize all types of information. Such readiness should include machine routines for conversion of all types of data to usable form.

The particulars of phase (2) involve a choice of a system of differential equations describing the motion of the satellite and the selection of a scheme for integrating them. The systems of equations fall into the following categories:

(1) Those involving elements of the osculating ellipse.

(2) Those involving coordinates of points on the orbit.

The formulation of the problem in terms of the equations of (1) above gives rise to the methods of Variation of Parameters discussed in Section I. The disadvantages of these methods are in general:

(1) Each method has exceptional cases which must be excluded.

(2) The equations are more complicated, and for a single integration time-step involve more extensive calculations which result in increased machine time and error generation.

(3) In Hansen's method, frequent use is made of series expansions, and hence the problem of "small divisors" may be introduced. This may cause large errors in computation.

(4) Encke's method requires frequent rectification of the ellipse, and hence may result in increased machine time.

The advantage of utilizing a set of differential equations of the first category lies in the fact that parameters may be chosen which vary slowly with respect to the independent variable. The error accumulation would be decreased.

In the second category, the main disadvantage is that the coordinates chosen may not vary as slowly, and thus make the error control more difficult.

The advantages of using a set of differential equations of the second category are:

(1) Direct approach to the problem which permits relatively simple equations. This leads to fewer computations and less error per integration step.

(2) There are no exceptional cases. This allows a single orbit computation program, which conserves personnel time and machine time.

(3) Reduction of machine time in conversion routines.

In either of the above categories, a selection of integration procedure must be made. The Runge-Kutta fourth order process has a truncation error of the order of $\underline{h}^5$ where $\underline{h}$ is the time-step of the integration. This requires that small time-steps be used which may result in an excessive round-off error. Provision may be made for increasing or decreasing the integration step-size. However, this is of questionable value, since an increase in step-size opens the way for large truncation error, and a decrease in step-size leads to greater round-off error. Furthermore, provision for changing the step-size may increase the machine running time.

Integration procedures using finite differences have the advantages of being relatively easy to apply, and the truncation error can be reduced considerably below that of the fourth order Runge-Kutta method, even with relatively low order differences. For example, with central differences, if n is the number of terms used in the formulae and errors of order higher than the first are neglected, then the order of differences used is $2n-2$ and the truncation error is of the order h^{2n+1} for time-step h. With forward or backward difference formulae, the truncation error is of the order of h^{2n}. Furthermore, for central differences only the even or the odd differences, not both, are present, whereas all differences occur in the forward or backward difference formulae. This implies that the central difference formulae are better than the forward or backward ones with regard to round-off error, since fewer machine operations are involved. The advantage of central differences over forward or backward differences is clear even when the difference table is discarded, and expressions derived in terms of preceding functional values. In these expressions the constant coefficients for the central differences are smaller numerically and alternate in algebraic sign. Hence the central difference formulae exhibit smaller oscillation amplitudes, and thereby converge more rapidly than the corresponding formulae using other differences. These finite difference techniques do have the minor objections of requiring auxiliary starting procedures for the integration and involving more work to change the interval of integration. However, for ephemeris computations it is not necessary to change the integration step-size. For example, if an observation is available at a time which does not coincide with a time corresponding to a step-time in the integration, a computed position at the time of the observation can be had to sufficient accuracy by use of an interpolation formula.

Regardless of the methods which are chosen for orbit work, the computation will ultimately deviate from true positions because of errors in the formulation of the equations of motion, the initial conditions, and computation procedures. Therefore, a third phase is necessary. This phase involves corrections on the orbit positions as calculated in phases (1) and (2). Differential correction formulae for various quantities describing the orbit in terms of elements of the osculating ellipse or the coordinates of positions on the orbit are known. It is recommended that corrections be applied directly to the rectangular coordinates of position. This is based upon the fact that explicit expressions can be obtained for the first order corrections. A differential correction scheme applied to rectangular coordinates is in the process of formulation, and will appear in another report.

In summary the recommendations are:

Phase (1): Since data may take many forms and arise from diverse sources, no one procedure for obtaining initial values can be recommended above all others. However, great care should be exercised in the selection of quantities to be observed or measured, and much thought devoted to the inherent errors. Furthermore, computation procedures converting from measured data to initial conditions must be considered thoroughly from the point of view of reducing errors in these initial conditions.

Phase (2):a) At present the best set of differential equations to be utilized is that giving the equations of motion in rectangular coordinates. The reasons for this are that no exceptional cases exist and the computation per integration step is minimized. b) Cowell's Method is recommended for use because of the simpler more direct calculations, the decreased truncation error, and the decreased round-off error. It is observed, however, that if particular cases should arise where Variation of Parameter methods are appropriate, central difference integration formulae should be applied.

Phase (3): First order differential correction procedures applicable to the rectangular coordinates of positions as calculated by Cowell's Method are recommended.

REFERENCES

1. Moulton, F. R., Celestial Mechanics. Second Revised Edition. New York, The Macmillan Company, 1923. 437 pp.

2. Herrick, S., Baker, R. M. L., and Hilton, C. G., "Gravitational and Related Constants for Accurate Space Navigation." Proceedings of the VIIIth International Astronautical Congress, Barcelona. pp. 199-235.1957.

3. Walters, L. G., and Herrick, S., "The Influence of the Earth's Potential Field on a Nearly Circular Satellite." Aeronutronic Systems, Inc., Publication, Newport Beach, California No. U-326. 1959.

4. Baker, R. M. L., "Application of Astronomical Perturbation Techniques to the Return of Space Vehicles." American Rocket Society Journal 29, No. 3, pp. 207-211. March 1959.

5. Sterne, T. E., Folkart, B. M., and Schilling, G. F., "An Interim Atmosphere Fitted to Preliminary Densities Inferred from U. S. S. R. Satellites." Smithsonian Institution Astrophysical Observatory, Cambridge, Research in Space Science, Special Report No. 21 pp. 1-37. 1959.

6. Porter, J. G., et. al. Planetary Coordinates. London, Her Majesty's Stationery Office. pp. 152-159. 1958. 160 pp.

7. Herget, P., and Musen, P., "A Modified Hansen Lunar Theory for Artificial Satellites." The Astronomical Journal 63, No. 10 pp. 430-433. November 1958 .

8. Herget, P., The Computation of Orbits. Privately Published 1948.

9. Herrick, S., Astrodynamics. (To be published by Von Nostrand in 1960.)

10. Jordan, C., Calculus of Finite Differences. New York. Chelsea Publishing Company. 1950. 652 pp.

11. Kopal, Z., Numerical Analysis. New York, John Wiley and Sons. 1955. 556 pp.

12. Kunz, K. S., Numerical Analysis. New York, McGraw-Hill Book Company. 1951. 381 pp.

13. Baker, R. M. L., et. al., "Efficient Precision Orbit Computation Techniques." American Rocket Society Publication. 1959. 40 pp.

14. Herrick, S., "Step-by-Step Integration of $\dot{x} = f(x,y,z,t)$ Without a Corrector." Mathematical Tables and Other Aids to Computation. No. 34 pp. 61-67. April 1951.

15. Cramer, H., Mathematical Methods of Statistics. Princeton, Princeton University Press, pp. 182. 1946. 575 pp.

16. Vienop, E., and Brady, J. L., "The Themis Code: An Astronomical Numerical Integration Program for the IBM 704." Livermore, California, University of California Radiation Laboratory Report 5242 pp. 20-27. 1958.

17. Sterne, T. E., "The Accuracy of Numerical Solutions of Ordinary Differential Equations." Mathematical Tables and Other Aids to Computation 7. No. 37-44. 1952,53. pp. 159-164.

APPENDIX A

GLOSSARY OF TERMS

Anomalistic period, T: The time between successive perigee passages of an orbiting object.

Argument of latitude, u: The instantaneous angle at the focus measured in the direction of motion of the object in the plane of its orbit, between the ascending node and the object.

Argument of perigee, ω: The angle at the focus measured in the direction of motion of the object in the plane of its orbit, between the ascending node and perigee.

Ascending node: The projection on the celestial sphere of the point where the object passes through the equatorial plane from south to north.

Eccentric anomaly, E: The angle measured from the center of the circle which circumscribes the orbit to the point on the circumcircle determined by the projection of a perpendicular to the semi-major axis passing through the instantaneous position of the object.

Eccentricity, e: A quantity, defined by the ratio of the center-to-focus distance to the semi-major axis, which measures the departure of an ellipse from its circumscribing circle.

Ecliptic: The projection on the celestial sphere of the apparent path of the sun.

Elements: Any set of independent parameters which uniquely define an osculating orbit.

Ephemeris: A tabular listing of the position of an object with respect to its reference frame.

Epoch: An arbitrary reference time.

Geoid: The spheroid described by some mathematical approximation to the actual shape of the earth. The geoid does not coincide with the surface of the earth except at one or more reference points.

Geocenter: The center of the geoid.

Heliocentric orbit: The orbit of a body moving under the predominant influence of the sun's gravitational attraction.

Inclination: The angle between the plane of the equator and the plane of the orbit.

Line of apsides: The line coinciding with the semi-major axis.

Longitude of ascending node, Ω: The angle at the focus measured eastward in the plane of the equator between the vernal equinox and the ascending node.

Longitude of object, u: The argument of latitude.

Longitude of perigee, π: The sum of the longitude of the ascending node and the argument of perigee.

Mean angular motion, n: The average angular rate at which the object moves in one complete circuit.

Mean anomaly, M: The product of the mean angular motion and the elapsed time from perigee.

Mean longitude, L: The sum of the mean anomaly and the longitude of perigee.

Parameter: One of a set of independent quantities which uniquely describe an orbit.

Perigee: The point on a geocentric orbit closest to the geocenter.

Residual: The difference between the observed location of the object and its location given by an ephemeris.

Semi-major axis, a: One-half of the longest diameter of an ellipse.

Sidereal angular velocity: The angular velocity of the earth with reference to the "fixed" stars.

True anomaly: The angle at the focus measured in the direction of motion of the object in its orbit plane between perigee and the instantaneous position of the object.

True longitude, ℓ: The sum of the true anomaly and the longitude of perigee.

Vernal Equinox: That point on the celestial sphere formed by the intersection of the projected ecliptic circle and the projected equatorial circle, where the apparent motion of the sun is from south to north.

APPENDIX B

TRANSFORMATION FROM GEOCENTRIC EQUATORIAL COORDINATES TO STANDARD ELEMENTS OF THE ELLIPSE

The equations for the transformation from the coordinates

$$x, \quad y, \quad z, \quad \dot{x}, \quad \dot{y}, \quad \dot{z}$$

to the standard elements of the ellipse

$$a, \quad e, \quad i, \quad \Omega, \quad \omega, \quad T_p$$

are given in the order used in computing, where

x, y, z are geocentric equatorial coordinates, with x and y lying in the equatorial plane, x pointing to the Vernal Equinox, y pointing 90° east of x, and z coinciding with the polar axes, positive north,

a is the semi-major axis of the ellipse,

e is the eccentricity,

i is the inclination angle, $0 \leq i \leq 180°$,

Ω is the longitude of the ascending node (right ascension of ascending node),

ω is the argument of perigee, and

T_p is the time of passage through perigee.

Obtain the radial distance to the object from geocenter from:

$$r = (x^2 + y^2 + z^2)^{1/2} \tag{1}$$

Obtain the square of the magnitude of the velocity from:

$$\dot{s}^2 = (\dot{x}^2 + \dot{y}^2 + \dot{z}^2) \tag{2}$$

Obtain the inner product of the radius and velocity vectors from:

$$\vec{r} \cdot \dot{\vec{r}} = x\dot{x} + \dot{y}y + z\dot{z} \tag{3}$$

Obtain the components of the vector outer product from:

$$\begin{aligned} h_x &= (y\dot{z} - z\dot{y}) \\ h_y &= -(x\dot{z} - z\dot{x}) \\ h_z &= (x\dot{y} - y\dot{x}) \end{aligned} \tag{4}$$

Obtain the semi-major axis from:

$$a = \frac{r}{2 - r\dot{s}^2} \tag{5}$$

Obtain the eccentricity from:

$$e = \left[(r\dot{s}^2 - 1)^2 + \left(\frac{\vec{r}\cdot\dot{\vec{r}}}{\sqrt{a}}\right)^2\right]^{1/2} \tag{6}$$

Obtain the eccentric anomaly from:

$$\tan E = \frac{\vec{r}\cdot\dot{\vec{r}}}{\sqrt{a}(r\dot{s}^2 - 1)} \tag{7}$$

with a quadrant check provided by:

$$e \cos E = (r\dot{s}^2 - 1) \tag{8}$$

Obtain the mean anomaly from:

$$M = E - e \sin E \tag{9}$$

Obtain the mean motion from:

$$n = a^{-3/2} \tag{10}$$

Obtain the time of perigee passage from:

$$T_p = t_o - \frac{M}{n} \tag{11}$$

Obtain the longitude of the ascending node from:

$$\tan \Omega = \frac{h_x}{-h_y} \tag{12}$$

with a quadrant check obtained from:

$$\sin i \cos \Omega = \frac{-h_y}{\sqrt{h_x^2 + h_y^2 + h_z^2}} \tag{13}$$

since sin i, by definition, is always positive.

Obtain the inclination angle from:

$$\tan i = \frac{\sqrt{h_y^2 + h_x^2}}{h_z} \tag{14}$$

The quadrant check is obtained from the fact that, by definition

$$0 \leq i \leq 180°$$

Obtain the true anomaly from:

$$v = 2 \tan^{-1} \sqrt{\frac{1 + e^2}{1 - e^2}} \tan E/2 \tag{15}$$

The argument of latitude is obtained from:

$$\tan u = \frac{z \sqrt{h_x^2 + h_y^2 + h_z^2}}{y h_x - x h_y} \tag{16}$$

with a quadrant check given by

$$r \sin i \cos u = \frac{y h_x - x h_y}{\sqrt{h_x^2 + h_y^2 + h_z^2}} \tag{17}$$

since r sin i is never negative.

The argument of perigee is obtained from:

$$\omega = u - v \tag{18}$$

The quantities have been obtained in the following order:

$$a, \quad e, \quad T_p, \quad \Omega, \quad i, \quad \omega$$

and the computation is complete.

The procedure given will fail in many places for zero eccentricity.

APPENDIX C

FIFTH DERIVATIVES, $x^{(v)}$, $y^{(v)}$, $z^{(v)}$, $u^{(v)}$, $v^{(v)}$, $w^{(v)}$

$$
\begin{aligned}
x^{(v)} = & \left\{\frac{a}{r^3} + \frac{abc}{r^5} + \frac{abz^2}{r^7}\right\} \ddddot{x} + \left\{\left(-\frac{9a}{r^4} - \frac{15abc}{r^6} - \frac{21abz^2}{r^8}\right)\dot{r} + \frac{6abz\dot{z}}{r^7}\right\} \dddot{x} \\
& + \left\{\left(-\frac{9a}{r^4} - \frac{15abc}{r^6} \quad \frac{21abz^2}{r^8}\right)\dddot{r} + \left(\frac{36a}{r^5} + \frac{90abc}{r^7} + \frac{168abz^2}{r^9}\right)(\dot{r})^2 - \frac{84abz\dot{z}\dot{r}}{r^8} + \frac{6abz\ddot{z}}{r^7} \right. \\
& \left. + \frac{6ab(\dot{z})^2}{r^7}\right\}\dot{x} + \left\{\left(-\frac{3a}{r^4} - \frac{5abc}{r^6} - \frac{7abz^2}{r^8}\right) \ddddot{r} + \left(\frac{36a}{r^5} + \frac{90abc}{r^7} + \frac{168abz^2}{r^9}\right)\dddot{r}\,\dot{r} \right. \\
& + \left(-\frac{60a}{r^6} - \frac{210abc}{r^8} - \frac{504abz^2}{r^{10}}\right)(\dot{r})^3 + \frac{2abz\ddddot{z}}{r^7} + \left(\frac{6ab\dot{z}}{r^7} - \frac{42abz\dot{r}}{r^8}\right)\ddot{z} + \left(-\frac{42abz\dddot{r}}{r^8}\right. \\
& \left. + \frac{336abz(\dot{r})^2}{r^9}\right)\dot{z} - \left.\frac{42ab\dot{r}(\dot{z})^2}{r^8}\right\} x + HV\ddddot{x} + 3H\dot{V}\dddot{x} + 3H\ddot{V}\ddot{x} + H\dddot{V}\dot{x} \; .
\end{aligned}
$$

$y^{(v)}$ may be obtained from $x^{(v)}$ by replacing x with y.

$$
\begin{aligned}
z^{(v)} = & \left(\frac{a}{r^3} + \frac{abc}{r^5} + \frac{3abz^2}{r^7}\right) \ddddot{z} + \left(-\frac{9a}{r^4} - \frac{15abc}{r^6} - \frac{63abz^2}{r^8}\right)\dddot{z}(\dot{r}) + \left(-\frac{9a}{r^4} - \frac{15abc}{r^6} - \frac{63abz^2}{r^8}\right)\dot{z} \\
& + \left(\frac{36a}{r^5} + \frac{90abc}{r^7} + \frac{504abz^2}{r^9}\right)\dot{z}(\dot{r})^2 + \left(-\frac{3az}{r^4} - \frac{5abcz}{r^6} - \frac{7abz^3}{r^8}\right) \ddddot{r} + \left(\frac{36az}{r^5} + \frac{90abcz}{r^7}\right. \\
& \left. + \frac{168abz^3}{r^9}\right)\dot{r}\,\dddot{r} + \left(-\frac{60az}{r^6} - \frac{210abcz}{r^8} - \frac{504abz^3}{r^{10}}\right)(\dot{r})^3 + \left(\frac{6ab}{r^7}\right)(\dot{z})^3 + \frac{18abz\dot{z}\dddot{z}}{r^7} \\
& + \left(-\frac{126abz}{r^8}\right)(\dot{z})^2\dot{r} + H V \overset{....}{z} + 3H\dot{V}\ddddot{z} + 3H\ddot{V}\dddot{z} + H\dddot{V}\dot{z}
\end{aligned}
$$

$$
\begin{aligned}
u^{(v)} = & \left\{\frac{a}{r^3} + \frac{abc}{r^5} + \frac{abz^2}{r^7}\right\} \overset{.....}{x} + \left\{\left(-\frac{12a}{r^4} - \frac{20abc}{r^6} - \frac{28abz^2}{r^8}\right)\dot{r} + \frac{8abz\dot{z}}{r^7}\right\} \ddddot{x} \\
& + \left\{\left(-\frac{18a}{r^4} - \frac{30abc}{r^6} - \frac{42abz^2}{r^8}\right)\ddot{r} + \left(\frac{72a}{r^5} + \frac{180abc}{r^7} + \frac{336abz^2}{r^9}\right)(\dot{r})^2 + \left(-\frac{168abz\dot{z}\dot{r}}{r^8}\right) \right. \\
& \left. + \frac{12abz\ddot{z}}{r^8} + \frac{12ab(\dot{z})^2}{r^7}\right\} \dddot{x} + \left\{\left(-\frac{12a}{r^4} - \frac{20abc}{r^6} - \frac{28abz^2}{r^8}\right) \ddddot{r} + \left(\frac{144a}{r^5} + \frac{360abc}{r^7}\right.\right. \\
& \left. + \frac{672abz^2}{r^9}\right) \dot{r}\,\dddot{r} + \left(-\frac{240a}{r^6} - \frac{840abc}{r^8} - \frac{2016abz^2}{r^{10}}\right)(\dot{r})^3 + \left(\frac{1344abz\dot{z}}{r^9}\right)(\dot{r})^2 +
\end{aligned}
$$

$$+ \left(- \frac{168ab(\dot{z})^2}{r^8} - \frac{168abz\ddot{z}}{r^8}\right)(\dot{r}) + \left(- \frac{168abz\dot{z}}{r^8}\right) \ddot{r} + \frac{24ab\dot{z}\ddot{z}}{r^7} + \frac{8abz\dddot{z}}{r^7} \Big\} \dot{x}$$

$$+ \Big\{\left(- \frac{3a}{r^4} - \frac{5abc}{r^6} - \frac{7abz^2}{r^8}\right) \ddddot{r} + \left(\frac{48a}{r^5} + \frac{120abc}{r^7} + \frac{224abz^2}{r^9}\right)\dot{r}\,\dddot{r} + \Big(- \frac{360a}{r^6}$$

$$- \frac{1260abc}{r^8} - \frac{3024abz^2}{r^{10}}\Big)(\dot{r})^2 \ddot{r} + \left(\frac{36a}{r^5} + \frac{90abc}{r^7} + \frac{168abz^2}{r^9}\right)(\ddot{r})^2 + \left(- \frac{56abz\dot{z}}{r^8}\right) \dddot{r}$$

$$+ \left(\frac{360a}{r^7} + \frac{1680abc}{r^9} + \frac{5040abz^2}{r^{11}}\right)(\dot{r})^4 + \left(\frac{1344abz\dot{z}\dot{r}\ddot{r}}{r^9}\right) + \frac{8ab\dot{z}\dddot{z}}{r^7} - \frac{1008abz\dot{z}(\dot{r})^3}{r^8}$$

$$- \frac{56abz\dot{r}\dddot{z}}{r^8} + \frac{2abz\ddddot{z}}{r^7} + \frac{672abz(\dot{r})^2\ddot{z}}{r^9} + \frac{6ab(\ddot{z})^2}{r^7} + \frac{672ab(\dot{r})^2(\dot{z})^2}{r^9} - \frac{168ab\dot{r}\dot{z}\ddot{z}}{r^8}$$

$$- \frac{3024abz\dot{z}(\dot{r})^3}{r^{10}} - \frac{84ab(\dot{z})^2\ddot{r}}{r^8} - \frac{84abz\ddot{r}\ddot{z}}{r^8} \Big\} x + 4H\dot{V}\ddddot{x} + HVx^{(v)} + 6H\ddot{V}\dddot{x} + 4H\dddot{V}\ddot{x}$$

$$+ H\ddddot{V}\dot{x} \quad .$$

$v^{(v)}$ may be obtained from $u^{(v)}$ by replacing x with y.

$$w^{(v)} = \left\{ \frac{a}{r^3} + \frac{abc}{r^5} + \frac{abz^2}{r^7} \right\} \ddddot{z} + \Big\{\left(- \frac{12a}{r^4} - \frac{15abd}{r^6} - \frac{5abc}{r^6} - \frac{70abz^2}{r^8}\right)\dot{r} + \frac{20abz\dot{z}}{r^7} \Big\} \dddot{z}$$

$$+ \Big\{\Big(- \frac{18a}{r^4} - \frac{30abd}{r^6} - \frac{126ab\ldots}{} \ldots [illegible] \ldots \frac{\ldots 8abz^2}{\ldots 9}\Big)(\dot{r})^2 - \frac{252abz\dot{z}\dot{r}}{r^8}$$

$$+ \frac{24ab(\dot{z})^2}{r^7} + \frac{18abz}{r^7} \ldots [illegible] \ldots + \Big(\frac{144a}{r^5} + \frac{360abd}{r^7}$$

$$+ \frac{1680abz^2}{r^9}\Big)\dot{r}\,\ddot{r} + \Big(- \ldots [illegible] \ldots \frac{\ldots abz\dot{z}\ddot{r}}{\ldots 8} + \frac{1008abz\dot{z}(\dot{r})^2}{r^9}$$

$$+ \frac{12ab\dot{z}\ddot{z}}{r^7} - \frac{168ab(\dot{z})^2}{r^8} \ldots [illegible] \ldots a + \frac{120abd}{r^7}$$

$$+ \frac{224abz^2}{r^9}\Big)\dot{r}\,\dddot{r} + \Big(- \frac{360a}{r^6} \ldots [illegible] \ldots \dot{z}\,\dddot{r} + \frac{336abz\dot{z}\dot{r}\ddot{r}}{r^9}$$

$$+ \Big(\frac{36a}{r^5} + \frac{90abd}{r^7} + \frac{168abz^2}{r^9} \ldots [illegible] \ldots \Big)^4 -$$

$$- \frac{1008ab z\dot{z}(\dot{r})^3}{r^{10}} + \frac{1008ab(\dot{z})^2(\dot{r})^2}{r^9} - \frac{252ab\dot{z}\,\dddot{z}\,\dot{r}}{r^8} - \frac{126ab(\dot{z})^2\ddot{r}}{r^8} \Big\} z$$

$$+ H V z^{(v)} + 4 H \dot{V} \ddddot{z} + 6 H \ddot{V} \dddot{z} + 4 H \dddot{V} \ddot{z} + H \ddddot{V} \dot{z} .$$

APPENDIX D

PREDICTOR AND CORRECTOR FORMULAE IN TERMS OF ORDINATES

The numerical integration formula

$$x_i = (\Delta t)^2 [''X_i + \frac{1}{12} X_i - \frac{1}{240} \delta^2 X_i + \frac{31}{60480} \delta^4 X_i - \frac{289}{3628800} \delta^6 X_i + \frac{317}{22809600} \delta^8 X_i - \frac{6803477}{2615348736000} \delta^{10} X_i] \tag{1}$$

which yields a predicted value for x_i in terms of entries in the difference table (4-18) can be converted into the double integration formula

$$\begin{aligned} x_i = (\Delta t)^2 [''X_i &+ 0.766936626\, X_{i-1} \\ &- 3.525812185\, X_{i-2} \\ &+ 10.1575288\, X_{i-3} \\ &- 19.87096838\, X_{i-4} \\ &+ 27.44087718\, X_{i-5} \\ &- 27.18762117\, X_{i-6} \\ &+ 19.29048178\, X_{i-7} \\ &- 9.59681704\, X_{i-8} \\ &+ 3.186400889\, X_{i-9} \\ &- 0.6352768226 X_{i-10} \\ &+ 0.05760362585 X_{i-11}] \end{aligned} \tag{2}$$

which expresses x_i in terms of $''X_i$, X_{i-1}, X_{i-2}, . . ., X_{i-11} .

Formula (2) is derived from formula (1) in the following manner: since the table of central differences (4-18) was extended from the table (4-13) so that all tenth order differences below $\delta^{10} X_{i-6}$ were equal to $\delta^{10} X_{i-6}$, it follows that

$$\delta^{10} X_i = \delta^{10} X_{i-1} = \delta^{10} X_{i-2} = \ldots = \delta^{10} X_{i-6} .$$

Furthermore, from the basic definitions of central differences, it follows that

$$\delta^{2K} X_j = \delta^{2K} X_{j-1} + \delta^{2K+1} X_{j-\frac{1}{2}} \qquad (3)$$

$$\delta^{2K+1} X_{j-\frac{1}{2}} = \delta^{2K+1} X_{j-\frac{3}{2}} + \delta^{2K+2} X_{j-1}$$

Now, then, from relation (3)

$$\delta^2 X_i = \delta^2 X_{i-1} + \delta^3 X_{i-\frac{1}{2}} \qquad (4)$$

From relation (4)

$$\delta^2 X_i = \delta^2 X_{i-1} + \delta^3 X_{i-\frac{3}{2}} + \delta^4 X_{i-1}$$

By continuing to apply relations (3) and (4), the following formula is obtained

$$\delta^2 X_i = \delta^2 X_{i-1} + \delta^3 X_{i-\frac{3}{2}} + \delta^4 X_{i-2} + \delta^5 X_{i-\frac{5}{2}} + \delta^6 X_{i-3} + \delta^7 X_{i-\frac{7}{6}} + \delta^8 X_{i-4} + \delta^9 X_{i-\frac{9}{2}} + \delta^{10} X_{i-5} . \qquad (5)$$

In a similar manner, formulae

$$\delta^2 X_{i-1} = \delta^2 X_{i-2} + \delta^3 X_{i-\frac{5}{2}} + \delta^4 X_{i-3} + \delta^5 X_{i-\frac{7}{2}} + \delta^6 X_{i-4} + \delta^7 X_{i-\frac{9}{2}} + \delta^8 X_{i-5} + \delta^9 X_{i-\frac{11}{2}} + \delta^{10} X_{i-6}$$

$$\delta^3 X_{i-\frac{3}{2}} = \delta^3 X_{i-\frac{5}{2}} + \delta^4 X_{i-3} + \delta^5 X_{i-\frac{7}{2}} + \delta^6 X_{i-4} + \delta^7 X_{i-\frac{9}{2}} + \delta^8 X_{i-5} + \delta^9 X_{i-\frac{11}{2}} + \delta^{10} X_{i-6}$$

$$\delta^4 X_{i-2} = \delta^4 X_{i-3} + \delta^5 X_{i-\frac{7}{2}} + \delta^6 X_{i-4} + \delta^7 X_{i-\frac{9}{2}} + \delta^8 X_{i-5} + \delta^9 X_{i-\frac{11}{2}} + \delta^{10} X_{i-6}$$

$$\delta^5 X_{i-\frac{5}{2}} = \delta^5 X_{i-\frac{7}{2}} + \delta^6 X_{i-4} + \delta^7 X_{i-\frac{9}{2}} + \delta^8 X_{i-5} + \delta^9 X_{i-\frac{11}{2}} + \delta^{10} X_{i-6}$$

$$\delta^6 X_{i-3} = \delta^6 X_{i-4} + \delta^7 X_{i-\frac{9}{2}} + \delta^8 X_{i-5} + \delta^9 X_{i-\frac{11}{2}} + \delta^{10} X_{i-6}$$

$$\delta^7 X_{i-\frac{7}{2}} = \delta^7 X_{i-\frac{9}{2}} + \delta^8 X_{i-5} + \delta^9 X_{i-\frac{11}{2}} + \delta^{10} X_{i-6}$$

$$\delta^8 X_{i-4} = \delta^8 X_{i-5} + \delta^9 X_{i-\frac{11}{2}} + \delta^{10} X_{i-6}$$

$$\delta^9_{\,i-\frac{9}{2}} = \delta^9 X_{i-\frac{11}{2}} + \delta^{10} X_{i-6}$$

$$\delta^{10}_{\,i-5} = \delta^{10} X_{i-6}$$

are obtained.

Substituting the above formulae in (5), it follows that

$$\delta^2 X_i = \delta^2 X_{i-2} + 2\delta^3 X_{i-\frac{5}{2}} + 3\delta^4 X_{i-3} + 4\delta^5 X_{i-\frac{7}{2}} + 5\delta^6 X_{i-4} + 6\delta^7 X_{i-\frac{9}{2}} + 7\delta^8 X_{i-5} + 8\delta^9 X_{i-\frac{11}{2}} + 9\delta^{10} X_{i-6}. \quad (6)$$

Now the central differences can be expressed in terms of $X_{i-1}, X_{i-2}, \ldots,$ X_{i-11} in the following manner:

$$\delta^2 X_{i-2} = X_{i-1} - 2X_{i-2} + X_{i-3},$$

$$2\delta^3 X_{i-\frac{5}{2}} = 2X_{i-1} - 6X_{i-2} + 6X_{i-3} - 2X_{i-4},$$

$$3\delta^4 X_{i-3} = 3X_{i-1} - 12X_{i-2} + 18X_{i-3} - 12X_{i-4} + 3X_{i-5},$$

$$4\delta^5 X_{i-\frac{7}{2}} = 4X_{i-1} - 20X_{i-2} + 40X_{i-3} - 40X_{i-4} + 20X_{i-5} - 4X_{i-6},$$

$$5\delta^6 X_{i-4} = 4X_{i-1} - 30X_{i-2} + 75X_{i-3} - 100X_{i-4} + 75X_{i-5} - 30X_{i-6} + 5X_{i-7},$$

$$6\delta^7 X_{i-\frac{9}{2}} = 6X_{i-1} - 42X_{i-2} + 126X_{i-3} - 210X_{i-4} + 210X_{i-5} - 126X_{i-6} + 42X_{i-7} - 6X_{i-8}$$

$$7\delta^8 X_{i-5} = 7X_{i-1} - 56X_{i-2} + 196X_{i-3} - 392X_{i-4} + 490X_{i-5} - 392X_{i-6} + 196X_{i-7} - 56X_{i-8} + 7X_{i-9},$$

$$8\delta^9 X_{i-\frac{11}{2}} = 8X_{i-1} - 72X_{i-2} + 288X_{i-3} - 672X_{i-4} + 1008X_{i-5} - 1008X_{i-6} + 672X_{i-7} - 288X_{i-8} + 72X_{i-9} - 8X_{i-10},$$

$$9\delta^{10} X_{i-6} = 9X_{i-1} - 90X_{i-2} + 405X_{i-3} - 1080X_{i-4} + 1890X_{i-5} - 2268X_{i-6} + 1890X_{i-7} - 1080X_{i-8} + 405X_{i-9} - 90X_{i-10} + 9X_{i-11}$$

Substituting these relations in formula (6), it follows that:

$$\delta^2 X_i = 45X_{i-1} - 330X_{i-2} + 1155X_{i-3} - 2508X_{i-4} + 3696X_{i-5} - 3828X_{i-6} + 2805_{i-7} - 1430X_{i-8} + 484X_{i-9} - 98X_{i-10} + 9X_{i-11}. \quad (7$$

The same procedure is followed in expressing $X_i; \delta^4 X_i, \ldots, \delta^{10} X_i$ in terms of $X_{i-1}, X_{i-2}, \ldots, X_{i-11}$. Substituting these expressions in formula (1), one obtains formula (2).

The double integration correction formula

$$
\begin{aligned}
x_i = (\Delta t)^2[''X_i &+ 0.05760362585 \quad X_i \\
&+ 0.1332967418 \quad X_{i-1} \\
&- 0.357612765 \quad X_{i-2} \\
&+ 0.652930536 \quad X_{i-3} \\
&- 0.861771849 \quad X_{i-4} \\
&+ 0.828002030 \quad X_{i-5} \\
&- 0.574745365 \quad X_{i-6} \\
&+ 0.2812852632 \quad X_{i-7} \\
&- 0.0922187772 \quad X_{i-8} \\
&+ 0.01820146864 \quad X_{i-9} \\
&- 0.001636938289 \quad X_{i-10} \,]
\end{aligned}
$$

can be obtained from formula (1) in a similar fashion. In this case the table of differences is extended so that

$$\delta^{10} X_i = \delta^{10} X_{i-1} = \ldots \delta^{10} X_{i-5} .$$

By utilizing formulae (3) and (4) and expressions for the central differences in terms of ordinates, $\delta^2 X_i$, $\delta^4 X_i, \ldots,$ and $\delta^{10} X_i$ can be expressed in terms of X_i, $X_{i-1}, \ldots, X_{i-10}$, which leads to formula (8).

In a similar manner, the single integration predictor and corrector formulae (4-21) and (4-23) can be obtained from

$$\dot{x}_i = \Delta t[\,'X_i - \frac{1}{12}\mu\delta X_i + \frac{11}{720}\mu\delta^3 X_i - \frac{191}{60480}\mu\delta^5 X_i + \frac{2497}{3628800}\mu\delta^7 X_i - \frac{14797}{95800320}\mu\delta^9 X_i]. \quad (9)$$

In this case, in addition to formùlae (3) and (4), use must be made of the formula $\mu\delta^{2K+1} X_i = \delta^{2K+1} X_{i-\frac{1}{2}} + \frac{1}{2}\delta^{2K+2} X_i$. The reduction follows the procedure given for deriving formula (2) from formula (1) .

It should be pointed out that many formulae similar to (2) can be derived from formula (1). These formulae would express x_i in terms other than the preceding X_j . The same statement applies to formula (8) and formulae (4-21) and (4-23) .

IV. FORMULAE EMPLOYING RECTANGULAR COORDINATES FOR THE DIFFERENTIAL CORRECTION OF ORBITS OF NEAR EARTH SATELLITES

1. Introduction

A differential correction procedure is a systematic method for using the residuals to: a) confirm that the set of orbit parameters is the best obtainable, or b) to adjust the orbit parameters to obtain improved values. The residuals are the differences between the actual observations and the values of the quantities observed computed from a set of orbit parameters.

A differential correction procedure provides the orbit analyst with a means for using experimental observations to gain an improved knowledge of the trajectory of a near earth satellite. This improved knowledge implies more than the ability to make more precise predictions: it also implies the knowledge of certain physical characteristics of the satellite, of the gravitational field, and of the atmosphere in which the satellite is moving. If these physical characteristics are not known, or are known imperfectly, they must be determined from experimental observations. At the present time, some form of differential correction procedure is the only means of utilizing widely separated observations. In addition, the so-called "irreducible residuals," which are residuals remaining after successive application of the correction procedure, are a sensitive diagnostic aid toward evaluating the performance of the various parts of a satellite tracking range-orbit computation complex. By examining these residuals, the orbit analyst can often separate and identify shortcomings in the theory, in the performance of individual stations, in station location surveys, etc.

Formulae for differentially correcting the orbit of a near earth satellite are given in this report. These formulae have been derived from an approximate solution to the equations of motion, including drag and oblateness effects. The approximate solution and the correction formulae are expressed in rectangular coordinates. Short period terms have been suppressed. Except for two simple formulae for evaluating average drag effects, numerical integrations are avoided.

The correction formulae are intended for use with a precise numerical integration procedure [1]* for computing the residuals, in which the integration is performed in rectangular coordinates. The values of a set of reference position and velocity coordinates are corrected.

A brief discussion of the concepts underlying the use of a differential correction procedure is given in sections 2 and 3. The approximate solution to the equations of motion in rectangular coordinates, and the derivation of analytical expressions for the partial derivatives are given in section 4. A complete tabulation of all formulae is given in section 5. Section 6 contains the summary.

*The numbers refer to the bibliography in Section 7.

2. Fundamental Concepts

Before developing the differential corrections, some of the fundamental concepts involved will be discussed.

Suppose that $v_1, v_2, \ldots, v_n$ are n observations of the quantity v. The true or exact value of v cannot be found from measurements due to the fact that the unit of measure and v are incommensurable. In addition, there are errors of some type in all measurements. Hence the error can never be found if the error is defined as the true value minus the measured value. However, something can be said regarding the magnitude of errors and their distribution. The function

$$G(v) = \frac{h}{\sqrt{\pi}} \varepsilon^{-h^2 v^2} \tag{2-1}$$

describes a normal distribution for h constant. The equation (2-1) has the interpretation that the relative number of observations whose errors lie within the limits v and dv is

$$\frac{h}{\sqrt{\pi}} \varepsilon^{-h^2 v^2} dv$$

It now follows that if the errors are $v - v_1$, $v - v_2$, $\ldots$, $v - v_n$, the probabilities that these individual errors will occur are:

$$P_1 = \frac{h}{\sqrt{\pi}} \varepsilon^{-h^2(v - v_1)^2} dv_1; \cdots; \quad P_n = \frac{h}{\sqrt{\pi}} \varepsilon^{-h^2(v - v_n)^2} dv_n$$

These are independent so the probability that these will occur simultaneously is

$$P = \left(\frac{h}{\sqrt{\pi}}\right)^n \varepsilon^{-h^2 \sum_{i=1}^{n} (v - v_i)^2} \prod_{i=1}^{n} dv_i$$

The objective is to choose v in such a way that this probability will be a maximum. This occurs when

$$\sum_{i=1}^{n} (v - v_i)^2$$

is a minimum, or when

$$\sum_{i=1}^{n} (v - v_i) = 0, \quad \text{or} \quad v = \frac{1}{n} \sum_{i=1}^{n} v_i \quad .$$

This states that the most probable value of v is the arithmetic mean of the observed values. For the case of n observations of equal reliability, the Principle of Least Squares follows:

> The best or most probable value obtainable from a set of n measurements, or observations of equal reliability, is that value for which the sum of the squares of the errors is a minimum.

For that case where all the measurements are not of equal reliability, the value of h above will be different for each of the observations and

$$P = \frac{\prod_{i=1}^{n} h_i}{(\sqrt{\pi})^n} \, \varepsilon^{-\sum_{i=1}^{n} [h_i(v - v_i)]^2} \prod_{i=1}^{n} dv_i$$

The best value of v occurs when $\sum_{i=1}^{n} [h_i(v - v_i)]^2$ is minimum. The introduction of the concept of weight w which means that an observation of weight w is equal in importance to w observations of unit weight, makes it possible to write, for $h_i^2 = w_i$ that P is maximum when $\sum_{i=1}^{n} w_i(v - v_i)^2$ is minimum. This gives the statement:

> The best value or most probable value obtainable from a set of n measurements, or observations of unequal reliability, is that value for which the sum of the weighted squares of the errors is minimum.

Since it is not possible to determine the errors, the residuals afford a method of prodedure. The difference between the most probable value of v and v_i is the residual for the i^{th} measurement. It can be shown [2] that the sum of the squares of the residuals is minimum when the sum of the squares of the errors is minimum and conversely. Therefore, in the preceding discussion the errors may be replaced by the residuals.

Suppose that a set of unknown quantities x, y, z, $\cdots$ are to be determined from n observations M_1, M_2, M_3, $\cdots$, M_n, where x, y, z, $\cdots$ are related by the linear relations, or equations of condition:

$$\begin{aligned} a_1x + b_1y + c_1z + \cdots &= M_1 \\ a_2x + b_2y + c_2z + \cdots &= M_2 \\ &\vdots \\ a_nx + b_ny + c_nz + \cdots &= M_n \end{aligned}$$

The a_i, b_i, c_i are assumed known and the M_i are subject to errors of an undetermined nature. Suppose that each of the M_i has a measure of reliability h_i, then the probability that the error Δ_i will occur in M_i is

$$\frac{h_i}{\sqrt{\pi}} \varepsilon^{-h_i^2 \Delta_i^2}$$

The probability that errors Δ_1, Δ_2, ..., Δ_n will occur is

$$P = \frac{\prod_{i=1}^{n} h_i}{(\sqrt{\pi})^n} \varepsilon^{-\sum_{i=1}^{n} (h_i \Delta_i)^n}$$

The most probable values of x, y, z, $\cdots$ are those for which this is a maximum. When the Δ_i are residuals, they are given by

$$\Delta_i = a_ix + b_iy + c_iz + \cdots - M_i \qquad (i = 1,2,\ldots, N) .$$

Hence, the most probable value of x, y, z, ... are those determined by $\sum_{i=1}^{n} (h_i\Delta_i)^2$ a minimum. For the h_i^2 replaced by the weight w_i, this becomes $\sum_{i=1}^{n} w_i\Delta_i^2$ a minimum.

It is convenient to introduce the bracket [ ] as a special summation symbol such that

$$[a\ a] = \sum_{i=1}^{n} a_i^2, \quad [a\ b] = \sum_{i=1}^{n} a_i\ b_i, \text{ etc.}$$

This permits the writing of the sum of w_i multiplied by the squares of Δ_i as:

$$[w\ a\ a]x^2 + [w\ b\ b]y^2 + [w\ c\ c]z^2 + \cdots$$

$$+ 2[w\ a\ b]x_y + \cdots + [w\ M\ M] .$$

A necessary condition that this be a minimum is that its partials with respect to the variables x, y, z, $\cdots$ be zero, or

$$\begin{array}{l} [w\ a\ a]x + [w\ b\ a]y + [w\ c\ a]z + \cdots = [w\ M\ a] \\ [w\ a\ b]x + [w\ b\ b]y + [w\ c\ b]z + \cdots = [w\ M\ b] \\ [w\ a\ c]x + [w\ b\ c]y + [w\ c\ c]z + \cdots = [w\ M\ c] \\ \vdots \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \end{array}$$

These are the normal equations. A simplification is achieved by multiplying each of the equations of condition by the square root of its weight. The normal equations now have the same weight.

A change in notation puts the above equations in the form:

$$\sum_{j=1}^{k} a_{ij}\ x_j = \beta_i \qquad (i = 1, 2, \cdots, n)$$

or in matrix form $AX = B$.

It is difficult to comment conclusively on the best way, for all cases, of solving a set of equations of the above form. Probably there is no method which can be called <u>best</u> for all types and solutions. However, it can be said that the best method to use on a particular set of equations must be dictated by the nature of the equations, the form of the solutions required, and the knowledge and experience of the computation analyst.

The properties of the equations that are of interest are:

a) whether or not the coefficients are small integers;

b) whether or not there are numerous zero coefficients;

c) whether or not the non-zero coefficients are systematically arranged;

d) whether or not the coefficients are exact or approximations involving measurement and rounding errors, and

e) the nature of the diagonal elements.

The following considerations concerned with the solution influence the choice of method:

a) whether or not a solution is required for one set or many sets of the β_i;

b) whether or not the properties of the matrix A are needed.

When all the foregoing have been carefully considered, a choice of either a direct method or an iterative scheme is made. When solutions are required for many sets of the β_i it may be best to evaluate A^{-1} and use the equation $x = A^{-1} B$. However, rounding errors may be involved in the calculations of A^{-1} and the solution at any one step would be regarded as approximations to the **x** .

Regardless of the choice of method, there are certain possible situations which must receive attention. If the determinant of the coefficients almost vanishes, numerical errors may produce an undetected nonsensical solution. More generally, if the equations possess the property of being ill-conditioned, or the determinant of the coefficients is small compared to any of the terms formed in the sum of the products of the elements, then extreme care must be exercised. In these cases the value of the determinant may change radically when a small change occurs in one or more of the coefficients. The ultimate solution may be very responsive to round-off errors or other errors in the coefficients. Hartree [3] cites the example:

$$x + 2y = 4$$

$$1000x + 2001y = 4003$$

for which $D = 1$, $x = -2$, $y = 3$. When the coefficient of y in the second equation is decreased by 0.1 per cent, or 1,999 the solution is $x = 10$, $y = -3$. An increase in the same coefficient of y of 0.1 per cent to 2,001 gives $x = 2$, $y = 1$.

The measure of ill-conditionedness of the equations may be obtained from the ratio of the maximum and minimum characteristic values of the matrix of the coefficients. If this ratio is near unity, the equations are well-conditioned; when the ratio is large compared to unity they are ill-conditioned. For other measures of condition see [4]. All of these tests are time consuming in their application, and are not too useful in practice. It is noted that another characteristic of ill-conditioned equations is that a set of values of the unknowns greatly different from the true solutions can give values for the β_i which differ from the true β_i by small amounts. This suggests that iterative tests of this type on the accuracy of the solution can fail. Finally, it can be shown [5] that if the original equations are ill-conditioned, then the normal equations will be even more ill-conditioned.

In view of the foregoing remarks it is evident that any method depending on matrix techniques for solving systems of linear equations may fail in isolated cases. However, the system arising in this problem contains only six equations which ordinarily possess few if any zero coefficients. Hence in general, Gaussian elimination which has a long and successful history of use [13], [14], [15], [16], [17] is recommended. In particular the Crout Modification which may be programmed directly from the detailed outline in [13] or [14] is recommended for this problem.

3. Differential Corrections

The basic concept behind the differential correction procedure can be illustrated as follows:

Let $M_i(x_o', y_o', \cdots, t_i)$ be an observation (e.g., a measurement of the local declination of the object) at time $t = t_i$, resulting from a set of unknown initial values $x_o', y_o', \cdots, \dot{z}_o'$ at $t = t_o$, and

$C_i(x_o, y_o, \cdots, t_i)$ be the computed value of the quantities observed at time $t = t_i$, based upon a set of known initial conditions $x_o, y_o, z_o, \cdots, \dot{z}_o$ at time $t = t_o$, where

$x_o, y_o, z_o, \dot{x}_o, \dot{y}_o, \dot{z}_o$ are the initial position and velocity coordinates.

If M_i is different from C_i in value, it is a function of a set of different initial values (neglecting, of course, observational and other errors). However, if M_i is not too far different from C_i, then M_i may be expanded in a Taylor's series about $x_o, y_o, \cdots, \dot{z}_o$

$$M_i = C_i(x_o, y_o, \cdots, \dot{z}_o, t_i) + \frac{\partial M_i}{\partial x_o}(x_o' - x_o) + \frac{\partial M_i}{\partial y_o}(y_o' - y_o) + \cdots \tag{3-1}$$

If the difference between the two sets of initial conditions is small, then the partial derivatives of the functions M_i may be replaced by the partial derivatives of the known functions C_i in (3-1),

$$\Delta\, \Theta_i \cong \frac{\partial C_i}{\partial x_o}\delta x_o + \frac{\partial C_i}{\partial y_o}\delta y_o + \cdots + \frac{\partial C_i}{\partial \dot{z}_o}\delta \dot{z}_o \tag{3-2}$$

where

$$\begin{aligned} \Delta\, \Theta_i &= M_i - C_i \equiv \text{the residual at } t = t_i \\ \delta x_o &= (x_o' - x_o) \\ \delta y_o &= (y_o' - y_o) \\ &\cdots\cdots\cdots \\ \delta \dot{z}_o &= (\dot{z}_o' - \dot{z}_o) \end{aligned}$$

The quantities $x_o', y_o', \cdots, \dot{z}_o'$ are the correct, unknown values of the initial conditions, which are to be found. If six observations are made (e.g., three observations of azimuth angles and three observations of elevation angle, or six observations of slant range, etc.) then (3-2) can be used to form a set of simultaneous equations in the unknowns δx_o, δy_o, $\cdots$, $\delta \dot{z}_o$

$$\begin{aligned}
\Delta\,\Theta_1 &= A_{11}\delta x_o + A_{12}\delta y_o + \cdots + A_{16}\delta \dot{z}_o \\
\Delta\,\Theta_2 &= A_{21}\delta x_o + A_{22}\delta y_o + \cdots + A_{26}\delta \dot{z}_o \\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
\Delta\,\Theta_6 &= A_{61}\delta x_o + A_{62}\delta y_o + \cdots + A_{66}\delta \dot{z}_o
\end{aligned} \tag{3-3}$$

where the A_{11}, A_{12}, $\cdots$, A_{ij} are the partial derivatives of the computed position with respect to each of the initial conditions, evaluated at time $t = t_1$, $t = t_2$, $\cdots$, $t = t_i$.

Solving (3-3) for δx_o, δy_o, $\cdots$, $\delta \dot{z}_o$, then the adjusted values of the initial conditions are given by

$$\begin{aligned}
x_o' &= x_o + \delta x_o \\
y_o' &= y_o + \delta x_o \\
&\cdots\cdots\cdots \\
\dot{z}_o' &= \dot{z}_o + \delta \dot{z}_o
\end{aligned} \tag{3-4}$$

From (3-1), (3-2), and (3-3), it is apparent that

a) a differential correction procedure is based on the <u>linearization</u> of a function relating the change in position at time $t = t_i$ to a change in position at time $t = t_o$,

b) from the conditions imposed in the derivation, there is no assurance that equations (3-3) will always have a solution, and

c) a rapid, accurate computation of the coefficients A_{11}, A_{12}, $\cdots$, A_{ij} is essential to the practical application of the procedure.

The remainder of this section is a discussion of these points.

The linearization resulting from the neglect of the squares and higher powers in (3-1) means that the solution to (3-3) will not yield the correct values for the initial conditions when substituted into (3-4). However, if the residuals are sufficiently small, the error will be negligible. Furthermore, it is easy to show from physical reasoning that, for small eccentricities, the higher derivatives in (3-1) do not increase in magnitude. Thus if the squares of the values of the quantities δx_o, δy_o, $\cdots$, $\delta \dot{z}_o$ are small, and equations (3-3) have a unique solution, the initial conditions as obtained from (3-4) are closer to the true values than the original set. These facts suggest an iterative procedure to obtain the initial conditions, using as a test for convergence some statistical measure of the magnitude of the residuals. This is the technique usually used in practice.

Equations (3-3) possess a unique solution if the determinant of the coefficients does not vanish, i.e.,

$$\begin{vmatrix} A_{11} & A_{12} & A_{13} & \cdots & A_{16} \\ A_{21} & A_{22} & A_{23} & \cdots & A_{26} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ A_{61} & A_{62} & A_{63} & \cdots & A_{66} \end{vmatrix} \neq 0$$

But the determinant of the coefficients is the Jacobian of the functions

$$\begin{aligned} M_1 &= M_1\,(x_o, y_o, z_o, \cdots, t_1) \\ M_2 &= M_2\,(x_o, y_o, z_o, \cdots, t_2) \\ &\cdots\cdots\cdots\cdots \\ M_6 &= M_6\,(x_o, y_o, z_o, \cdots, t_6) \end{aligned} \tag{3-5}$$

with respect to the parameters x_o, y_o, z_o, $\cdots$, $\dot{z}_o$. A non-vanishing Jacobian insures the existence of unique inverses to functions (3-5) of the form

$$
\begin{aligned}
x_o &= f_1\,(M_1, M_2, \cdots, M_6) \\
y_o &= f_2\,(M_1, M_2, \cdots, M_6) \\
&\cdots\cdots\cdots\cdots \\
\dot{z}_o &= f_6\,(M_1, M_2, \cdots, M_6)
\end{aligned}
\qquad (3\text{-}6)
$$

Thus, the conditions for a unique solution to equations (3-3) are that the observations contain enough information to yield a unique, independent set of parameters x_o, y_o, z_o, $\cdots$, $\dot{z}_o$ at $t = t_o$. For example, if equations (3-3) are set up for observations concentrated exactly at perigee, say, then the equations (3-3) will not have a unique solution. A test for this condition can be had from the test for a non-vanishing determinant. The difficulty here is that if the determinant "almost vanishes," numerical errors may produce a nonsensical solution that will go undetected. Rather than to rely entirely on the test of the determinant, it is better to make sure that the observations are well distributed around the orbit. To increase the probability that the observations are well distributed (and, incidentally, to reduce the effects of observational errors, as discussed in section 1) the set of equations (3-3) is usually replaced by an overdetermined set of equations. If a "least squares" criteria is applied to the "fit" of the trajectory to the observations, then one form of the overdetermined set is given by the six normal equations:

$$
\begin{aligned}
\Big(\sum_{i=1}^{n} \Delta\,\Theta_i A_{i1}\Big) &= \Big(\sum_{i=1}^{n} A_{i1}A_{i1}\Big)\delta x_o + \Big(\sum_{i=1}^{n} A_{i2}A_{i1}\Big)\delta y_o + \cdots \Big(\sum_{i=1}^{n} A_{i6}A_{i1}\Big)\delta \dot{z}_o \\
\Big(\sum_{i=1}^{n} \Delta\,\Theta_i A_{i2}\Big) &= \Big(\sum_{i=1}^{n} A_{i1}A_{i2}\Big)\delta x_o + \Big(\sum_{i=1}^{n} A_{i2}A_{i2}\Big)\delta y_o + \cdots \Big(\sum_{i=1}^{n} A_{i6}A_{i2}\Big)\delta \dot{z}_o \\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
\Big(\sum_{i=1}^{n} \Delta\,\Theta_i A_{i6}\Big) &= \Big(\sum_{i=1}^{n} A_{i1}A_{i6}\Big)\delta x_o + \Big(\sum_{i=1}^{n} A_{i2}A_{i6}\Big)\delta y_o + \cdots \Big(\sum_{i=1}^{n} A_{i6}A_{i6}\Big)\delta \dot{z}_o
\end{aligned}
\qquad (3\text{-}7)
$$

where n is the total number of observations. The solution of (3-7) is a least squares estimate of the correct solution based on n observations.*

The replacement of equations (3-3) by an overdetermined set is no guarantee that the solution will always exist, particularly if the satellite tracking complex is limited to a few stations. If a large number of stations are providing observations the probability of a solution is increased by using more observations in the equations. Some provision for preliminary editing prior to, and for close monitoring during the computational phase is necessary.

Two methods for determining A_{11}, A_{12}, $\cdots$, A_{ij} are

a) the variant method, in which the equations of motion are integrated numerically with small variations imposed on each coordinate in succession. The coefficients are determined from one of the approximate formulae for the partial derivative, for example, by

$$A_{i1} = \frac{C_i\,(x_o + \Delta x_o, y_o, z_o, \cdots, \dot{z}_o) - C_i\,(x_o, y_o, \cdots, \dot{z}_o)}{\Delta x_o} \qquad (3\text{-}8)$$

where $C_i\,(x_o + \Delta x_o, y_o, \cdots)$ is the numerical integration of the equations of motion from $t = t_o$ to $t = t_i$ with a variation Δx_o imposed on the coordinate x_o.

b) The derivation of an analytical expression for the partial derivatives, obtained from an approximate solution to equations of motion. (If drag, oblateness, and other perturbing effects are negligible, an exact solution is available, of course.)

The variant method requires a greater amount of computing time and, because of the accumulation of rounding and truncation error, probably yields results no more accurate than those obtained from the approximate analytical solution

*The form of equations (3-7) requires the accumulation of sums which may grow large enough, with large n, to cause an overflow in the computer. This is particularly true with radar data, because of the high rate of data acquisition possible, and the usual reluctance to discard all but a few points during a passage of the satellite. With optical data, however, the possibility of having an excessive number of observations in a given period of time apparently is not so great.

to the equations of motion. To illustrate, if equations (3-8) are used, for six observations, 37 separate numerical integrations are required. If an over-determined set of equations is used, six additional integrations are required for each additional observation. Equations (3-8) are the simplest, and the least accurate, form of the approximate equations for the derivative. If more precise formulae are used, the number of numerical integrations required is at least doubled. Rounding and truncation error may be reduced by going to more complex, higher order integration schemes, but at the expense of additional computing time. It is apparent that the computing time requirements may easily become excessive when the variant method is used.

Therefore, the use of the analytical expressions for the partial derivatives results in a saving in computing time requirements and in increased accuracy.

It is recommended that the analytical expressions for the partial derivatives be used whenever practicable.

The derivation of the coefficients A_{12}, A_{13}, ..., A_{ij} is given in the next section.

4. Approximations of the Partial Derivatives

A. The Differential Equations of Motion: Approximations

The derivation of the analytical forms for the partial derivatives is based on an approximate solution to the equations of motion. The equations of motion, in rectangular coordinates, are given by

$$\begin{aligned} \ddot{x} &= -\frac{x}{r^3} + F_x \\ \ddot{y} &= -\frac{y}{r^3} + F_y \\ \ddot{z} &= -\frac{z}{r^3} + F_z \end{aligned} \qquad (4\text{-}1)$$

where x, y, z are the geocentric equatorial rectangular coordinates,

r is the distance from geocenter to satellite,

F_x, F_y, F_z are forces acting upon the satellite other than the forces due to the attraction of a perfectly spherical earth.

Equations (4-1) are expressed in canonical units, in which the equatorial radius of the earth is taken as the basic unit of length, and the unit of time is the time required for a hypothetical earth satellite, moving in a circular orbit with a semi-major axis equal to one equatorial earth radii, to traverse an arc distance of one radian.

The perturbing forces F_x, F_y, F_z are produced by a number of sources. For purposes here, only those forces due to oblateness and to the earth's atmosphere are considered. From reference [6], the gravitational potential due to the oblateness of the earth is given by

$$\Phi = \frac{J}{3r^3}\left[1 - \frac{3z^2}{r^2}\right] + \frac{H}{5r^4}\left[3\left(\frac{z}{r}\right) - 5\left(\frac{z}{r}\right)^3\right] + \frac{K}{30r^5}\left[3 - 30\left(\frac{z}{r}\right)^2 + 35\left(\frac{z}{r}\right)^4\right] \qquad (4\text{-}2)$$

where J, H, K are the second, third and fourth harmonics of the spherical harmonic expansion of the earth's gravitational field. The other terms have been defined previously. According to reference [7], the harmonic coefficients have the magnitudes

$$J = (1623.41 \pm 4) \times 10^{-6}$$
$$H = (6.04 \pm 0.73) \times 10^{-6}$$
$$K = (6.37 \pm 0.23) \times 10^{-6}$$

where the equatorial earth radius is taken to be 6378145($\pm$1) meters, and the canonical time unit is taken to be 806.811 seconds.

The J term contributes the major effect. To the order of terms of 10^{-6}, the H and K terms may be neglected.

The drag forces arising from motion through the atmosphere are given by

$$\begin{aligned} F_{x_D} &= -\frac{1}{2}\left(\frac{SC_D}{m}\right)\rho V(\dot{x} + \omega_e y) \\ F_{y_D} &= -\frac{1}{2}\left(\frac{SC_D}{m}\right)\rho V(\dot{y} - \omega_e x) \\ F_{z_D} &= -\frac{1}{2}\left(\frac{SC_D}{m}\right)\rho V \dot{z} \end{aligned} \tag{4-3}$$

where

S = the frontal area of the satellite

m = the mass of the satellite

C_D = the drag coefficient

ρ = the density of the atmosphere at satellite altitude

V = the magnitude of the velocity in geocentric equatorial coordinates

ω_e = the rotational velocity of the earth.

The other symbols have been defined previously. In canonical time units, the various terms in (4-3) have the order of magnitude

$$\frac{1}{2}\left(\frac{SC_D}{m}\right)\rho \sim 10^{-3} \text{ [e.g., 1958 } \varepsilon \text{ at 220 kilometers]}$$

$$V \sim 1$$

$$x,\ y \sim 1$$

$$\dot{x},\ \dot{y} \sim 1$$

$$\omega_e x,\ \omega_e y \sim 6 \times 10^{-5}$$

Therefore, the drag may contribute a total acceleration of the same order as the oblateness term. The term arising from the rotation of the atmosphere contributes an acceleration of the order of 10^{-5}. In view of the uncertainties associated with the drag force, for the purposes here it appears that it is better to omit this term.

With these approximations, the differential equations of motion are

$$\begin{aligned}
\ddot{x} &= -\frac{x}{r^3} - \frac{Jx}{r^5}\left[1 + 5\left(\frac{z}{r}\right)^2\right] - \frac{1}{2}\left(\frac{C_D S}{m}\right)\rho V \dot{x} \\
\ddot{y} &= -\frac{y}{r^3} - \frac{Jy}{r^5}\left[1 + 5\left(\frac{z}{r}\right)^2\right] - \frac{1}{2}\left(\frac{C_D S}{m}\right)\rho V \dot{y} \qquad (4\text{-}4) \\
\ddot{z} &= -\frac{z}{r^3} - \frac{Jz}{r^5}\left[3 + 5\left(\frac{z}{r}\right)^2\right] - \frac{1}{2}\left(\frac{C_D S}{m}\right)\rho V \dot{z}
\end{aligned}$$

B. The Approximate Solution

The drag and oblateness effects may be separated, since the drag forces act primarily to oppose the motion in the plane, whereas the oblateness forces act primarily to change the orientation of the plane. Cross coupling effects appear, but these are small.

The drag force causes a secular change in the semi-major axis and in the eccentricity of the instantaneous osculating ellipse. These may be determined from the changes in total energy and total angular momentum. The total energy may be obtained from (4-4) by multiplying the first, second, and third equations by dx, dy, and dz respectively, adding, and integrating. The result

$$\frac{1}{2a} - \frac{1}{2a_o} = \frac{J}{3r^3}\left[1 - 3\left(\frac{z}{r}\right)^2\right] - \frac{1}{2}\left(\frac{C_D S}{m}\right)\int_{t_o}^{t} \rho\, V^3\, dt \qquad (4\text{-}5)$$

is the equation for the semi-major axis of the instantaneous ellipse. The constant a_o is obtained from the energy integral

$$V^2 = \frac{2}{r} - \frac{1}{a}$$

for the two-body problem. The total angular momentum is related by a constant (the cosine of the inclination angle) to the angular momentum about the z axis, h, neglecting short period terms. The angular momentum about the z axis is obtained from (4-4) by multiplying the second equation by x, the first by y, and then subtracting the first from the second. The result

$$\frac{dh}{dt} + \frac{1}{2}\left(\frac{C_D S}{m}\right)\rho\, V\, h = 0 \qquad (4\text{-}6)$$

is the differential equation for the change in the total momentum. Equations (4-5) and (4-6) may be combined to obtain the secular changes in the semi-major axis and the eccentricity of the instantaneous oscillating ellipse. Sterne (Ref. [8]) has given the formulae for these changes. However, he recommends the use of alternate formulae giving the changes in perigee and apogee, namely,

$$\frac{dr_p}{dt} = -\left(\frac{C_D S}{m}\right)\sqrt{a}\,\frac{(1 - e)}{\pi}\int_o^{\pi}\left(\frac{1 + e\cos E}{1 - e\cos E}\right)^{\frac{1}{2}}(1 - \cos E)\,\rho\, dE$$

$$\frac{dr_a}{dt} = -\left(\frac{C_D S}{m}\right)\sqrt{a}\,\frac{(1 + e)}{\pi}\int_o^{\pi}\left(\frac{1 + e\cos E}{1 - e\cos E}\right)^{\frac{1}{2}}(1 + \cos E)\,\rho\, dE \qquad (4\text{-}7)$$

on the basis that a rough numerical integration of the integrals in (4-7) will yield a higher accuracy than a similar integration of the equations for a and e. In equations (4-7)

r_p is the perigee distance,

r_a is the apogee distance,

E is the eccentric anomaly,

and the other symbols have been defined.

Equations (4-7) give the average change in the apogee and perigee heights per revolution. Periodic variations have been suppressed. These formulae were obtained from a "variation of arbitrary constants" solution to the equations of motion [8]; they may also be obtained from the equations given by Moulton, Reference [9], page 405.

The semi-major axis and the eccentricity are obtained from (4-7) by

$$a = \frac{r_a + r_p}{2}$$
$$e = \frac{r_a - r_p}{r_a + r_p} \tag{4-8}$$

The major effect of the oblateness is a secular rotation of the line of nodes and the line of apsides. The mean angular motions are given by

$$\frac{d\Omega}{dt} = - \frac{Jn}{p^{5/2}} h$$
$$\frac{d\omega}{dt} = - \frac{2Jn}{p^2} [4 - 5 \frac{h^2}{p}] \tag{4-9}$$

where

Ω is the longitude of the ascending node,

ω is the argument of perigee,

$\sqrt{p}$ is the total angular momentum and

h is the angular momentum about the z-axis .

From (4-9) and the two-body solution, a solution to (4-4) in terms of the initial coordinates can be constructed. Define the following

$$\begin{aligned}
\Delta\Omega &= -\frac{Jhn}{p^{5/2}}\tau \\
\Delta\omega &= -\frac{2J}{p^2}\left[4 - \frac{5h^2}{p}\right] n\tau \\
F &= \cos\Delta\omega - \left(\frac{r\dot{r}}{\sqrt{p}}\right)\sin\Delta\omega \\
F' &= \frac{r^2}{\sqrt{p}}\sin\Delta\omega \\
G &= \cos\Delta\Omega \\
G' &= \sin\Delta\Omega \\
f &= 1 - \frac{a}{r_o}(1-\cos\Delta E) \\
f' &= \frac{a^2 n\cos\Delta E}{rr_o} \\
g' &= 1 - \frac{a}{r}(1-\cos\Delta E) \\
g &= \tau - \frac{(\Delta E - \sin\Delta E)}{n}
\end{aligned} \tag{4-10}$$

where

$\tau = t - t_o$, in canonical time units,

E is the eccentric anomaly,

$\Delta E = E - E_o$, where

E_o is the eccentric anomaly at $t = t_o$

r_o is the radial distance at $t = t_o$.

n is the mean angular motion per unit time.

The other symbols have been defined. The functions f, g, f', g' are given in Reference [10] page 48, and appear in the two-body solution. The solution to (4-4) to the order of terms in 10^{-6}, is given by

$$x = [F \cdot f + F' \cdot f'] [Gx_o - G'y_o] + [F \cdot g + F' \cdot g'] [G\dot{x}_o - G'\dot{y}_o]$$

$$y = [F \cdot f + F' \cdot f'] [Gy_o + G'x_o] + [F \cdot g + F' \cdot g'] [G\dot{y}_o + G'\dot{x}_o] \quad (4\text{-}11)$$

$$z = [F \cdot f + F' \cdot f'] z_o + [F \cdot g + F' \cdot g'] \dot{z}_o$$

where x_o, y_o, z_o, $\dot{x}_o$, $\dot{y}_o$, $\dot{z}_o$ are the coordinates at $t = t_o$.

The short period terms in (4-11) have been suppressed. It is emphasized that (4-11) has been derived for the purpose of obtaining coefficients in the differential correction equations. It is not intended that (4-11) be used to compute an ephemeris.

C. The Total Differentials

The derivatives of (4-11) with respect to x_o, y_o, z_o, ..., $\dot{z}_o$ are to be determined.

From (4-10)

$$df = f\left(\frac{dr_o}{r_o} - \frac{da}{a}\right) - \frac{rf'}{an} d(\Delta E)$$

$$dg = -\frac{3}{2}(g - \tau)\left(\frac{da}{a}\right) + (1-f)\frac{r_o}{a} d(\Delta E)$$

$$df' = f'\left(\frac{1}{2}\frac{da}{a} + \frac{1}{r_o}\frac{dr_o}{r_o}\right) - \frac{af'}{r}\cos E \cdot de + \frac{\sqrt{a}}{rr_o}\left[1 + \frac{r}{a}(g'-1)\right] d(\Delta E)$$

$$dF = -\left[\sin\Delta\omega + \frac{r\dot{r}}{\sqrt{p}}\cos\Delta\omega\right]d(\Delta\omega) - \sin\Delta\omega\left[\frac{1 + e^2}{(1 - e^2)^2}\sin E\right] de$$

$$dF' = \frac{r^2}{\sqrt{p}}\cos\Delta\omega\, d(\Delta\omega) + \sin\Delta\omega\, d\left(\frac{r^2}{\sqrt{p}}\right) \quad (4\text{-}12)$$

$$dG = -\sin\Delta\Omega\, d(\Delta\Omega)$$

$$dG' = \cos\Delta\Omega\, d(\Delta\Omega)$$

The manipulations required to reduce (4-12) into terms of dx_o, dy_o, ..., $d\dot{z}_o$ are rather tedious. The many relationships, both geometrical and dynamical, available from the two-body solution are essential to the reduction. The following is given to illustrate the procedure.

To shorten the notation, let X indicate the positions, Y the velocities, and let the superscript i signify a summation, i.e.,

$$dx_o + dy_o + dz_o = (dX)^i$$

$$d\dot{x}_o + d\dot{y}_o + d\dot{z}_o = (dY)^i$$

etc.

From the two-body solution

$$a = \frac{r_o}{2 - r_o G^2}$$

and

$$G^2 = \dot{x}_o^2 + \dot{y}_o^2 + \dot{z}_o^2 \tag{4-13}$$

$$\left(\frac{da}{a}\right) = \frac{2a}{r_o^3}(X\,dX)^i + 2a\,(Y\,dY)^i$$

is obtained. Therefore

$$\left(\frac{da}{a} - \frac{dr_o}{r_o}\right) = \frac{aG^2}{r_o^2}(X\,dX)^i + 2a\,(Y\,dY)^i$$

$$\left(\frac{1}{2}\frac{da}{a} + \frac{dr_o}{r_o^2}\right) = \left(\frac{a+1}{r_o^3}\right)(X\,dX)^i + a\,(Y\,dY)^i\,. \tag{4-14}$$

From the total differentials of

$$e \cos E_o = \left(1 - \frac{r_o}{a}\right)$$

$$e \sin E_o = \left(\frac{r\dot{r}}{\sqrt{a}}\right)_o \tag{4-15}$$

the total differential de is obtained

$$de = \frac{r_o}{a}\cos E_o\left(\frac{da}{a} - \frac{dr_o}{r_o}\right) - \frac{\sin E_o}{\sqrt{a}}\left[(X\,dY - Y\,dX)^i - a\,(Y\,dY + X\,dX)^i\right] \tag{4-16}$$

The differential $d(\Delta E)$ may be obtained from (4-15), but the result will contain the eccentricity in the denominator, so that for quasi-circular orbits the formula may fail. It is better to use instead

$$n(t - t_o) = \Delta E - e(\sin E - \sin E_o)$$

from which

$$d(\Delta E) = -\frac{3}{2}\frac{nr_o}{a}\left(\frac{da}{a}\right)\cdot\tau + \sin E \cdot de \qquad (4\text{-}17)$$

The final form for $d(\Delta E)$ is obtained with the aid of (4-13), (4-14), and (4-16).

The procedure is continued along the lines indicated. The end products of the manipulations are the partial derivatives which appear in the equations

$$\begin{aligned} dx &= \left(\frac{\partial x}{\partial x_o}\right) dx_o + \left(\frac{\partial x}{\partial y_o}\right) dy_o + \left(\frac{\partial x}{dz_o}\right) dz_o \cdots + \left(\frac{\partial x}{\partial \dot{z}_o}\right) d\dot{z}_o \\ dy &= \left(\frac{\partial y}{\partial x_o}\right) dx_o + \left(\frac{\partial x}{\partial y_o}\right) dy_o + \quad \cdots \quad + \left(\frac{\partial y}{\partial \dot{z}_o}\right) d\dot{z}_o \\ dz &= \left(\frac{\partial z}{\partial x_o}\right) dx_o + \left(\frac{\partial z}{\partial y_o}\right) dy_o + \quad \cdots \quad + \left(\frac{\partial z}{\partial \dot{z}_o}\right) d\dot{z}_o \end{aligned} \qquad (4\text{-}18)$$

The partial derivatives are given in Chart I as coefficients of dx_o, dy_o, ..., $d\dot{z}_o$.

D. The Transformation to the Observer's Coordinates

The total differentials in (4-18) must be expressed in terms of the observed quantities. The following forms of the observations are considered:

1) local right ascension and declination,

2) local azimuth and elevation,

3) slant range .

For the purpose here, it will be assumed that the usual corrections (e.g., corrections to the elevation angle for the deviation of the vertical, etc.) have been made to the data.

The coordinates of the satellite are

$$
\begin{aligned}
\rho \cos\theta \cos\beta &= x - X \\
\rho \sin\theta \cos\beta &= y - Y \\
\rho \cos\beta &= z - Z
\end{aligned}
\tag{4-19}
$$

where

ρ is the slant range from observer to satellite,

θ is the local right ascension of the satellite,

β is the local declination of the satellite,

x, y, z are the geocentric equatorial coordinates of the satellite, and

X, Y, Z are the geocentric equatorial coordinates of the observer.

From (4-19)

$$
\begin{aligned}
\rho \cos\beta \, d\theta &= -\sin\theta \cdot dx + \cos\theta \cdot dy \\
\rho \, d\beta &= -\cos\theta \sin\beta \cdot dx - \sin\theta \sin\beta \cdot dy + \cos\beta \cdot dz \\
d\rho &= \cos\theta \cos\beta \cdot dx + \sin\theta \cos\beta \cdot dy + \sin\beta \cdot dz
\end{aligned}
\tag{4-20}
$$

is obtained by forming the total differentials of both sides of (4-19), and solving. The location of the observer is assumed to be known. (4-20) expresses the residuals in the observed quantities directly in terms of the variations in the reference coordinates δx_o, δy_o, $\cdots$, $\delta \dot{z}_o$ by means of (4-18) and Table I. In the first two equations of (4-20), the value of ρ on the left hand side must be the computed value if the data has been obtained optically. The third equation is to be used with slant range data (e.g., radar range).

The angle residuals on the left hand side of (4-20) may be computed directly from the rectangular coordinates and the observed angles by means of the equations

$$
\begin{aligned}
\rho \cos\beta \cdot d\theta &= (x - X) \sin\theta - (y - Y) \cos\theta \\
\rho \cdot d\beta &= [(x - X)^2 + (y - Y)^2]^{1/2} \sin\beta - (z - Z) \cos\beta
\end{aligned}
\tag{4-21}
$$

where the rectangular coordinates are the computed values at the time of the observation (i.e., obtained from an integration of the equations of motion) and the angles are the observed values. Equation (4-21) may be verified by substituting (4-19) into the right hand side, introducing the trigonometric identities, and using the assumption that the angle residual is small.

If the observed angles are azimuth and elevation, defined as the angle measured east from the observer's true north point to the satellite in the observer's horizon plane, and the angle measured upward from the observer's horizon plan to the object, then an additional transformation is required. The local right ascension and declination are obtained from the transformation

$$\cos\beta\cos\theta = -\cos\Gamma_o\sin\varphi_o\cos\alpha\cos\delta - \sin\Gamma_o\sin\alpha\cos\delta + \cos\Gamma_o\cos\varphi_o\sin\delta$$

$$\cos\beta\sin\theta = -\sin\Gamma_o\sin\varphi_o\cos\alpha\cos\delta + \cos\Gamma_o\sin\alpha\cos\delta + \sin\Gamma_o\cos\varphi_o\sin\delta$$

$$\sin\beta = \cos\varphi_o\cos\alpha\cos\delta + \sin\varphi_o\sin\delta \tag{4-22}$$

where

δ is the elevation angle,

α is the azimuth angle,

Γ_o is the right ascension of the observer, and

φ_o is the declination of the observer.

The subtleties of locating the observer in geocentric coordinates and of locating the observer's horizon plane, which require definitions of the figure of the earth and of time, have been avoided. For a treatment of these problems, see Reference [11].

5. Tabulation of Formulae

The coefficients in the differential correction formulae are given in Chart I. The coefficients are written in a form to emphasize the symmetry.

The formulae are long, and, at first sight, complex. However, certain terms recur frequently. Because of the recurrence of many terms, it does not appear that machine computation of the coefficients will present a problem from the viewpoint of storage locations or from the viewpoint of machine computing time.

CHART I.

THE COEFFICIENTS IN THE DIFFERENTIAL CORRECTION EQUATIONS*

$\frac{\partial x}{\partial x_o}$	$G\sigma + \psi \frac{x_o}{r_o^2} - J\{(\bar{\lambda}\sigma + \bar{\mu}\epsilon)[\frac{x_o}{r_o^2}(N + \frac{2a}{r_o}M) + \frac{n\dot{y}_o t}{p^{5/2}}] - (f\lambda + g\mu)(\frac{ILx_o}{r_o^2} + \frac{20C\dot{y}_o}{\sqrt{p}})t - (f'\lambda + g'\mu)[\frac{x_o}{r^2}(\Lambda + \frac{2a}{r_o}\Gamma) + \frac{20C\dot{y}_o t}{\sqrt{p}}\cos\Delta\omega]\}$
$\frac{\partial x}{\partial y_o}$	$-G'\sigma + \psi \frac{y_o}{r_o^2} - J\{(\bar{\lambda}\sigma + \bar{\mu}\epsilon)[\frac{y_o}{r_o^2}(N + \frac{2a}{r_o}M) - \frac{n\dot{x}_o t}{p^{5/2}}] - (f\lambda + g\mu)(\frac{ILy_o}{r_o^2} - \frac{20C\dot{x}_o}{\sqrt{p}})t - (f'\lambda + g'\mu)[\frac{y_o}{r_o^2}(\Lambda + \frac{2a}{r_o}\Gamma) - \frac{20C\dot{x}_o t}{\sqrt{p}}\cos\Delta\omega]\}$
$\frac{\partial x}{\partial z_o}$	$+\psi \frac{z_o}{r_o^2} - J\{(\bar{\lambda}\sigma + \bar{\mu}\epsilon)[\frac{z_o}{r_o^2}(N + \frac{2a}{r_o}M)] - (f\lambda + g\mu)(\frac{ILz_o}{r_o^2}) - (f'\lambda + g'\mu)[\frac{z_o}{r_o^2}(\Lambda + \frac{2a}{r_o}\Gamma)]\}$
$\frac{\partial x}{\partial \dot{x}_o}$	$G\epsilon + \psi'\dot{x}_o - J\{(f\lambda + g\mu)(\frac{20Cy_o t}{\sqrt{p}}) + (f'\lambda + g'\mu)(\frac{20Cy_o t}{\sqrt{p}}\cos\Delta\omega) + (\bar{\lambda}\sigma + \bar{\mu}\epsilon)(2a\dot{x}_o M - \frac{ny_o t}{p^{5/2}})\}$
$\frac{\partial x}{\partial \dot{y}_o}$	$-G'\epsilon + \psi'\dot{y}_o - J\{(f\lambda + g\mu)(\frac{20Cx_o t}{\sqrt{p}}) + (f'\lambda + g'\mu)(\frac{20Cx_o t}{\sqrt{p}}\cos\Delta\omega) + (\bar{\lambda}\sigma + \bar{\mu}\epsilon)(2a\dot{y}_o M - \frac{nx_o t}{p^{5/2}})\}$
$\frac{\partial x}{\partial \dot{z}_o}$	$\psi'\dot{z}_o - J\{(\bar{\lambda}\sigma + \bar{\mu}\epsilon)(2aM)\dot{z}_o\}$

* The symbols are defined in terms of previously defined functions at the end of the Table.

CHART I (continued).

THE COEFFICIENTS IN THE DIFFERENTIAL CORRECTION EQUATIONS

$\dfrac{\partial y}{\partial x_o}$	$G'\sigma + \bar{\psi}\,\dfrac{x_o}{r_o^{\,2}} - J\{(\lambda\sigma + \mu\epsilon)\,[\dfrac{x_o}{r_o^{\,2}}\,(N + \dfrac{2a}{r_o}M) + \dfrac{n\dot{y}_o t}{p^{5/2}}] - (f\bar{\lambda} + g\bar{\mu})\,(\dfrac{ILx_o}{r_o^{\,2}} + \dfrac{20C\dot{y}_o}{\sqrt{p}})t - (f'\bar{\lambda} + g'\bar{\mu})\,[\dfrac{x_o}{r^2}\,(\Lambda + \dfrac{2a}{r_o}\Gamma) + \dfrac{20C\dot{y}_o t}{\sqrt{p}}\cos\Delta\,\omega]\}$
$\dfrac{\partial y}{\partial y_o}$	$+ G\sigma + \bar{\psi}\,\dfrac{y_o}{r_o^{\,2}} - J\{(\lambda\sigma + \mu\epsilon)\,[\dfrac{y_o}{r_o^{\,2}}\,(N + \dfrac{2a}{r_o}M) - \dfrac{n\dot{x}_o t}{p^{5/2}}] - (f\bar{\lambda} + g\bar{\mu})\,(\dfrac{ILy_o}{r_o^{\,2}} - \dfrac{20C\dot{x}_o}{\sqrt{p}})t - (f'\bar{\lambda} + g'\bar{\mu})\,[\dfrac{y_o}{r_o^{\,2}}\,(\Lambda + \dfrac{2a}{r_o}\Gamma) - \dfrac{20C\dot{x}_o t}{\sqrt{p}}\cos\Delta\,\omega]\}$
$\dfrac{\partial y}{\partial z_o}$	$+ \bar{\psi}\,\dfrac{z_o}{r_o^{\,2}} - J\{(\lambda\sigma + \mu\epsilon)\,[\dfrac{z_o}{r_o^{\,2}}\,(N + \dfrac{2a}{r_o}M) - (f\bar{\lambda} + g\bar{\mu})\,(\dfrac{ILz_o}{r_o^{\,2}}) - (f'\bar{\lambda} + g'\bar{\mu})\,[\dfrac{z_o}{r_o^{\,2}}\,(\Lambda + \dfrac{2a}{r_o}\Gamma)]\}$
$\dfrac{\partial y}{\partial \dot{x}_o}$	$G'\epsilon + \bar{\psi}'\dot{x}_o - J\{(f\bar{\lambda} + g\bar{\mu})\,(\dfrac{20Cy_o t}{\sqrt{p}}) + (f'\bar{\lambda} + g'\bar{\mu})\,(\dfrac{20Cy_o t}{\sqrt{p}}\cos\Delta\,\omega) + (\lambda\sigma + \mu\epsilon)\,(2a\dot{x}_o M - \dfrac{ny_o t}{p^{5/2}})\}$
$\dfrac{\partial y}{\partial \dot{y}_o}$	$+ G\epsilon + \bar{\psi}'\dot{y}_o - J\{(f\bar{\lambda} + g\bar{\mu})\,(\dfrac{20Cx_o t}{\sqrt{p}}) + (f'\bar{\lambda} + g'\bar{\mu})\,(\dfrac{20Cy_o t}{\sqrt{p}}\cos\Delta\,\omega) + (\lambda\sigma + \mu\epsilon)\,(2a\dot{y}_o M - \dfrac{n_o x_o t}{p^{5/2}})\}$
$\dfrac{\partial y}{\partial \dot{z}_o}$	$\bar{\psi}'\dot{z}_o - J\{(\lambda\sigma + \mu\epsilon)\,(2aM)\dot{z}_o\}$

CHART I (continued).

THE COEFFICIENTS IN THE DIFFERENTIAL CORRECTION EQUATIONS

$\frac{\partial z}{\partial x_o}$	$\Phi \frac{x_o}{r_o^2} + J\{(fz_o + g\dot{z}_o)(\frac{ILx_o t}{r_o^2} + \frac{20C\dot{y}_o t}{\sqrt{p}}) + (f'z_o + g'\dot{z}_o)(\frac{20C\dot{y}_o t}{\sqrt{p}} \cos \Delta\omega)\}$
$\frac{\partial z}{\partial y_o}$	$\Phi \frac{y_o}{r_o^2} + J\{(fz_o + g\dot{z}_o)(\frac{ILy_o t}{r_o^2} + \frac{20C\dot{x}_o t}{\sqrt{p}}) + (f'z_o + g'\dot{z}_o)(\frac{20\dot{x}_o t}{\sqrt{p}} \cos \Delta\omega)\}$
$\frac{\partial z}{\partial z_o}$	$\sigma + \Phi \frac{z_o}{r_o^2} + J\{(fz_o + g\dot{z}_o)(\frac{I\,Lz_o t}{r_o^2})\}$
$\frac{\partial z}{\partial \dot{x}_o}$	$\Phi' \dot{x}_o - J\{(fz_o + g\dot{z}_o) + (f'z_o + g'\dot{z}_o)\cos \Delta\omega\}\{\frac{20Cy_o t}{\sqrt{p}}\}$
$\frac{\partial z}{\partial \dot{y}_o}$	$\Phi' \dot{y}_o - J\{(fz_o + g\dot{z}_o) + (f'z_o + g'\dot{z}_o)\cos \Delta\omega\}\{\frac{20Cx_o t}{\sqrt{p}}\}$
$\frac{\partial z}{\partial \dot{z}_o}$	$\epsilon + \Phi' \dot{z}_o$

CHART I (continued).

THE COEFFICIENTS IN THE DIFFERENTIAL CORRECTION EQUATIONS

(a) Definition of the intermediate quantities:

$$A = -\frac{3}{2}\frac{nr_o}{a}$$

$$B = \left(\frac{16n}{p^2} - \frac{30h^2n}{p^3}\right)$$

$$C = \frac{hn}{p^{5/2}}$$

$$\xi = -\frac{3}{2}(g - t)$$

$$\zeta = (1 - f)\frac{r_o}{a}$$

$$\nu = (g' - 1)$$

$$\alpha = \frac{1}{\sqrt{a}\ r_o}\left[\frac{a}{r} + \nu\right]$$

$$\beta = \frac{r_o \cos E_o}{a(1 - \sin E \sin E_o)}$$

$$\gamma = \frac{2\beta ae}{p}$$

$$\eta = A\,t + \beta \sin E$$

$$P = \xi + \zeta\,\eta$$

$$Q = \zeta\,\beta \sin E$$

$$R = f - \sqrt{a}\ r\,\beta \sin E$$

$$S = f - \frac{rr_o}{a^2} f' t + \sqrt{a}\ r\,\beta \sin E$$

$$T = -f'\left(\frac{1}{2} + \frac{a}{r}\beta \cos E\right) + \alpha\,\eta$$

$$U = f'\left(\frac{1}{2} + \frac{a}{r}\beta \cos E\right) + \alpha\,\beta \sin E$$

$$V = \frac{a}{r}\nu\,\beta \cos E + \frac{r_o f'}{an}\beta \sin E$$

$$I = -\left(\sin \Delta\omega + \frac{rr}{\sqrt{p}} \cos \Delta\omega\right)$$

$$K = B\left(\frac{3}{2} - \gamma\right) - \frac{20n}{p^2}$$

$$L = B\,\gamma$$

$$M = C\left[4 - \frac{5}{2}\gamma p\right]$$

$$N = \frac{5}{2}C\gamma$$

CHART I (continued).

THE COEFFICIENTS IN THE DIFFERENTIAL CORRECTION EQUATIONS

(a) Definition of the intermediate quantities (continued):

$$\theta = JIKt - \frac{\beta(1+e^2)}{(1-e^2)^2} \sin E \, (\sin \Delta\,\omega)$$

$$\Gamma = JK\,(\cos \Delta\,\omega)t - \frac{r}{\sqrt{p}}\left(\frac{3}{2}\,r + \frac{\gamma}{4} - 2a\,\beta \cos E\right)$$

$$\Lambda = JN\,(\cos \Delta\,\omega)t - \frac{r}{\sqrt{p}}\left(\frac{r\gamma}{4} - 2\,a\,\beta \cos E\right)$$

$$\lambda = (Gx_o - G'y_o)$$

$$\bar{\lambda} = (Gy_o + G'x_o)$$

$$\mu = (G\dot{x}_o - G'\dot{y}_o)$$

$$\bar{\mu} = (G\dot{y}_o + G'\dot{x}_o)$$

$$\sigma = (F\,f + F'\,f')$$

$$\epsilon = (F\,g + F'\,g')$$

$$e \cos E_o = \left(1 - \frac{r_o}{a}\right)$$

$$e \sin E_o = \left(\frac{r\dot{r}}{\sqrt{a}}\right)_o$$

$$n = a^{-3/2}$$

$$f = 1 - \frac{a}{r_o}\,[1 - \cos(E - E_o)]$$

$$f' = \frac{a^2 n}{r r_o} \cos(E - E_o)$$

$$g = t - \left[\frac{(E - E_o) - \sin(E - E_o)}{n}\right]$$

$$g' = 1 - \frac{a}{r}\,[1 - \cos(E - E_o)]$$

$$F = \cos \Delta\,\omega - \left(\frac{r\dot{r}}{\sqrt{p}}\right) \sin \Delta\,\omega$$

$$F' = \frac{r^2}{\sqrt{p}} \sin \Delta\,\omega$$

$$G = \cos \Delta\,\Omega$$

$$G' = \sin \Delta\,\Omega$$

$$\psi = (f\lambda + g\mu)\left(\frac{2a\theta}{r_o}\right) + (f'\lambda + g'\mu)\left(\frac{2a\Gamma}{r_o} + \Lambda\right) + \lambda\,[F(R - \frac{2a}{r_o}S) + F'(U + \frac{2a}{r_o}T) - \mu\,[F(Q - \frac{2a}{r_o}\,P) - F'(W + \frac{2a}{r_o}V)]$$

$$\bar{\psi} = (f\bar{\lambda} + g\bar{\mu})\left(\frac{2a\theta}{r_o}\right) + (f'\bar{\lambda} + g'\bar{\mu})\left(\frac{2a\Gamma}{r_o} + \Lambda\right) + \bar{\lambda}[F(R - \frac{2a}{r_o}S) + F'(U + \frac{2a}{r_o}T) - \bar{\mu}\,[F\,(Q - \frac{2a}{r_o}P) - F'(W + \frac{2a}{r_o}V)]$$

$$\psi' = (f\lambda + g\mu)\,(2a\theta) + (f'\lambda + g'\mu)\,(2a\Gamma) + 2a\lambda\,[F'T - FS] + 2a\mu\,[F'V + FP]$$

$$\bar{\psi}' = (f\bar{\lambda} + g\bar{\mu})\,(2a\theta) + (f'\bar{\lambda} + g'\bar{\mu})\,(2a\Gamma) + 2a\bar{\lambda}\,[F'T - FS] + 2a\bar{\mu}\,[F'V + FP]$$

$$\Phi = (fz_o + g\dot{z}_o)\left(\frac{2a\theta}{r_o}\right) + (f'z_o + g'\dot{z}_o)\left(\frac{2a\Gamma}{r_o} + \Lambda\right) + z_o\,[F\,(R - \frac{2a}{r_o}S) + F'(U + \frac{2aT}{r_o})] - \dot{z}_o\,[F\,(Q - \frac{2aP}{r_o}) - F'(W + \frac{2a}{r_o}V)]$$

$$\Phi' = (fz_o + g\dot{z}_o)\,(2a\theta) + (f'z_o + g'\dot{z}_o)\,(2a\Gamma) + 2az_o\,[F'T - FS] + 2a\dot{z}_o\,[F'V + FP]$$

6. Summary

Formulae for differentially correcting the orbit of a near earth satellite are given. The formulae are based on an approximate solution, in rectangular coordinates, of the differential equations of motion. Drag and oblateness perturbations are included. The rectangular coordinates of velocity and position at a reference time are adjusted.

The procedure for using the formulae is outlined:

a) The residuals are computed from (4-20), using the observations and the rectangular coordinates obtained from an integration of the equations of motion. If the rectangular coordinates are not given at the time of the observation, the correct values may be obtained by interpolation. (The results of the interpolation may be made as exact as the tabulated values.)

b) The perigee and apogee altitudes at the time of the observation are computed from (4-7), using one of the standard model atmospheres (e.g., ARDC Model Atmosphere 1959). The integration of (4-7) is accomplished by simple quadratures, (e.g., Gauss's Numbers, Ref. 12). The constants of integration are the values of the perigee and apogee altitudes at the reference time, as computed from the reference coordinates.

c) The semi-major axis and the eccentricity are obtained from (4-8). From these, the component of angular momentum normal to the plane, and the mean motion are determined. The component of the momentum about the z-axis is determined from the values of the reference coordinates. The coefficients are computed, using (4-10) and Table I.

d) The procedure is repeated for n observations. The total time interval is limited primarily by the magnitude of the drag force, and the uncertainties in the drag force. It is believed that ten to twelve days will represent a reasonable limitation on the time interval.

e) Six simultaneous equations are formed from the n observations by (4-18) and (4-7). The equations are solved by one of the standard methods (e.g., Crout's Method) for the changes in the reference coordinates.

f) With new values of the reference coordinates, a new ephemeris and new residuals are computed.

7. References

1. Currie, J. C., et al. "Study of Satellite Orbit Computation Methods and an Error Analysis of the Mathematical Procedures," Technical Report No. 3, prepared under Department of the Air Force Contract No. AF 29(600)-1756, Project No. 1770, Headquarters Air Force Missile Development Center, Air Research and Development Command, Holloman Air Force Base, New Mexico, August 1959. 70 p.

2. Scarborough, J. B., Numerical Mathematical Analysis. Second Edition. Baltimore, The Johns Hopkins Press. 1950. 405 p.

3. Hartree, D. R., Numerical Analysis. Second Edition. Oxford, The Clarendon Press. 1958. 171 p.

4. Turing, A. M., Quarterly Journal of Mechanics and Applied Mathematics, (1) p. 287. 1948.

5. Taussky, Olga, Mathematical Tables and Other Aids to Computation. IV, Number 30, p. 111-112. April, 1950.

6. Walters, L. G. and Herrick, S., "The Influence of the Earth's Potential Field on a Nearly Circular Satellite," Aeronutronic Systems, Inc., Publication, Newport Beach, California, No. U 326. 1959.

7. Baker, R. M. L., Jr., et al., "Interim Report Covering Task Five, Range Planning Study for Project Mercury," Aeronutronic Systems, Inc., Publication, Newport Beach California, No. U 538. July 30, 1959.

8. Sterne, T. E., "An Atmospheric Model, and Some Remarks on the Inference of Density from the Orbit of a Close Earth Satellite," The Astronomical Journal, V. 63, No. 3, p. 81-87. March, 1958.

9. Moulton, F. R., Celestial Mechanics. Second Revised Edition. New York, The Macmillan Company, 1923. 437 p.

10. Herget, P., The Computation of Orbits. Privately Published 1948.

11. Veis, George, "Geodetic Applications of Observations of the Moon, Artificial Satellites and Rockets." Dissertation Presented in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy. The Ohio State University. 1958.

12. Milne, W. E., Numerical Calculus. Princeton, Princeton University Press. 1949.

13. Kunz, Kaiser S. Numerical Analysis. New York, McGraw-Hill Book Company, Inc. 1957. p. 226-229.

14. Scanlon, Robert H., and Rosenbaum, Robert. Aircraft Vibration and Flutter. New York, The MacMillan Company. 1951. p. 34-36.

15. Forsythe, George E., and Rosenbloom, Paul C. Numerical Analysis and Partial Differential Equations. New York, John Wiley and Sons, Inc. 1958. p. 21-27

16. Forsythe, George E. Bulletin of American Mathematical Society. Vol. 59, No. 4, p. 299-329. July, 1953.

17. Faddeeva, V. N. Computational Methods of Linear Algebra. New York, Dover Publications, Inc. p. 63-79.